HYPERSPECTRAL TARGET DETECTION PERFORMANCE MODELING

Thesis

Submitted to

The School of Engineering of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree of

Master of Science in Electrical Engineering

By

Christopher Joseph Morman

UNIVERSITY OF DAYTON

Dayton, Ohio

December, 2015

HYPERSPECTRAL TARGET DETECTION PERFORMANCE MODELING

Name: Morman, Christopher Joseph

APPROVED BY:

Russell Hardie, Ph.D.
Advisor Committee Chairman
Professor
Department of Electrical and Computer Engineering

Joseph Meola, Ph.D.
Committee Member
Electrical Engineer
Air Force Research Laboratory

Eric Balster, Ph.D.
Committee Member
Assistant Professor
Department of Electrical and Computer Engineering

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P.E.
Dean
School of Engineering

ABSTRACT

HYPERSPECTRAL TARGET DETECTION PERFORMANCE MODELING

Name: Morman, Christopher Joseph
University of Dayton

Advisor: Dr. Russell Hardie

Hyperspectral remote sensing has become a popular topic of research due to the numerous applications stemming from the high dimensionality of the recorded spectral data. From the design perspective, hyperspectral sensors are generally more complex than standard color or infrared imaging systems because there are more optical components in the system. The quality of each of these components directly affects the target detection performance of the system. In addition to the integrity of optical components, target detection performance is also affected by signal variations due to sensor noise. This research addresses the design of an end-to-end hyperspectral imaging system performance model that incorporates the optical design of the system as well as the stochastic nature of data collected by electronic remote sensing.

A system transmission model is presented that calculates the camera signal as a function of input radiance and accounts for each individual optical element in the imaging system. This model can be used to analyze the performance sensitivities of a specific component for a variety of target detection scenarios. The accuracy of the system transmission model is assessed using calibrated hyperspectral data. In addition to the system transmission model, a realistic statistical data model is proposed. Many data models currently account for sensor noise with an additive, stationary

variance. This research expands upon this by implementing an additive, signal-dependent sensor noise model that more accurately represents the true phenomena driving the sensor noise. The same data set is used to test target detection performance using the signal-dependent noise model. The results are analyzed to investigate the possible benefits of using the proposed noise model.

The data used for this research was collected at Wright Patterson Air Force Base 25-26 June 2014. The scene consists of a grassy background with eight painted wooden panel targets. Data collections took place at different times of day in order to capture varying solar angles and illumination levels. Additionally, data was collected with varying exposure times in an effort to observe performance effects due to varying signal-to-noise ratios. Conclusions about the performance of the system transmission and data modeling techniques are framed within the context of collection time and exposure time.

ACKNOWLEDGMENTS

Many people are deserving of recognition for their role in the completion of this research. I would like to thank Russ Hardie for acting as my advisor at the University of Dayton. Ever since my undergraduate years, I have been inspired by his enthusiasm for the field of image processing. I would also like to extend my thanks for Joe Meola for all of his help with my research and for serving on my committee. Whether I needed help understanding technical material or just wanted to swing by for a chat, Joe never hesitated to take time out of his day to help. I also thank Eric Balster for taking the time out of his busy schedule to serve on my committee.

I received much assistance from the staff in the Electrical Engineering department at the University of Dayton, and for that I am very grateful. Nancy Striebich worked with me every step of the way to ensure I stayed on top of all of my administrative requirements. From University of Dayton Research Institute, I'd like to acknowledge and thank Tom Fitzgerald for providing valuable feedback and support throughout the lifetime of this research.

My friends have provided me with incredible support for many years for which I am thankful. Finally, I would like to extend my thanks to my family. For years you have been my source of support and inspiration in all that I do. I could not have done this without you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

This chapter familiarizes the reader with several hyperspectral remote sensing concepts referenced throughout the document. It reviews previous hyperspectral performance modeling work and several common target detection algorithms, establishing the framework of this research. The system transmission and data models proposed by this research are introduced here as well. The chapter concludes with an outline of the remainder of the report.

1.1 Hyperspectral Remote Sensing

The field of hyperspectral remote sensing is a practical application of the combination of spatial imagery and molecular spectroscopy. As such, hyperspectral imaging (HSI) systems have the ability to capture information pertaining to the chemical makeup of an object from a remote location. Hyperspectral systems often operate between the visible and long-wave infrared portions of the electromagnetic spectrum (0.4-14 μm). The system used for this research operates in the visible to near-infrared (VNIR) range of 0.4-1.0 μm. In this region the primary source of electromagnetic radiation is reflective solar light. During data collection periods *spectral signatures* are formed by measuring the power fluctuations of this light at numerous contiguous narrow spectral bands. These spectral signatures can be used to detect and identify materials present within a pixel. Target detection algorithms look to take advantage of this spectral information to detect the presence of

a known material within a background (non-target) [1]. The performance of these algorithms is measured by the *probability of detection* (pD) – the probability that a pixel consisting of the target material is correctly identified as a target – and *false alarm rate* (FAR) – the probability that a background pixel is falsely identified as a target. A well designed target detection algorithm will have a high pD and a low FAR. In an effort to support HSI system design, Kerekes and Baum developed an end-to-end model that predicts system performance in terms of pD and FAR [2]. Their model can be broken into three distinct elements: HSI system modeling, data modeling, and atmospheric modeling. This research aims to expand upon the HSI system and data modeling functionality of Kerekes and Baum's model while leaving the development of advanced atmospheric modeling techniques for future work.

HSI system modeling consists of predicting the raw signal produced by the system as a function of a known illumination level. Kerekes and Baum use HSI systems that are well known throughout the hyperspectral community such as the airborne Hyperspectral Digital Imagery Collection Experiment (HYDICE) and Hyperion instruments. The performance characteristics of these sensors are well documented [3, 4, 5]. The model created by Kerekes and Baum allows for some parameter changes such as the number of spectral channels, transmission of fore-optics, and integration time, but the sensor system transmission is treated as a black box. This model does not allow for the analysis necessary to understand the relative importance of each optical element used in the system design.

A visual of a generic Offner dispersive spectrometer design in Figure 1.1 shows a number of the optical components within an HSI system. This example system consists of a lens, three mirrors, a diffraction grating, and a focal plane array (FPA). Each component has a spectrally dependent efficiency that factors into the overall system transmission. This research aims to expand upon the HSI system model used by Kerekes and Baum by implementing a system model that accounts

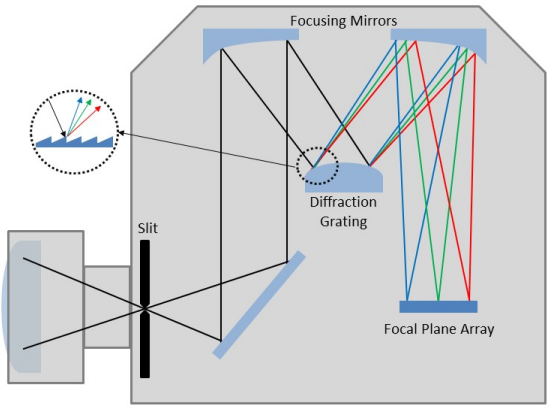


Figure 1.1: Offner Spectrometer Design. This figure provides a visual describing the design of the HSI system used for this research. Light enters through a vertical slit and is focused onto a reflective diffraction grating which reflects the light at different angles depending on wavelength. This light is then reflected once more and focused onto the focal plane array.

for each individual element's contribution to the system transmission. If the specifications of each individual element are known it should follow that a mathematical transfer function exists that is able to fully model the data collection process. This design would allow for the model to investigate, for example, which of several proposed mirror designs will result in the highest level of performance for a given target and background. Another advantage of this model is that it would allow near instantaneous performance comparisons between proposed HSI system designs.

While a simple transfer function is expected to describe the system model, a stochastic approach must be applied for data modeling. Kerekes and Baum assume that classes of interest can be represented by first and second-order spectral statistics – an assumption that is shared with this research. Additionally, they model sensor noise with a constant variance that is comprised of a summation of multiple system parameters. However, true hyperspectral noise exhibits signal-dependent characteristics that their model does not address. The noise in hyperspectral imagery is often classified as either fixed pattern noise or random noise. This work assumes the fixed pattern noise is made negligible by radiometric calibration. According to remote sensing theory, the remaining random

noise is the result of two independent random processes – circuitry noise and photon noise [6]. Theory dictates that the circuitry noise is signal-independent and Normally distributed while the photon noise is signal-dependent and Poisson distributed. However, a Normal approximation of the Poisson process may be used to model this signal-dependent component while operating in a photon rich environment [7]. This research attempts to improve upon Kerekes and Baum's noise model by using a more accurate noise model that takes all of these considerations into effect. The result is an additive signal-dependent noise model specific to hyperspectral imagery that was proposed by Meola et. al. [8]. This improved model allows for this research to investigate the relative importance of noise model accuracy in predicting target detection performance.

In the field of hyperspectral target detection, a majority of algorithms are developed from statistical data models in which background statistics are treated as stationary. In real data, background classes often display spectral variability due to reflectance variations within the class among other physical processes. For instance, consider a grassy scene. Even though everything is of the same class – grass – the visible colors may range from brown to green. Additionally, the grass has texture and depth that affect its appearance. In addition to this background clutter, due to the electronic data collection process, real data also exhibits variability due to signal-dependent sensor noise. Many algorithms, such as the adaptive cosine estimator (ACE) [1] and spectral matched filter (SMF) [9] model the data variability using a stationary mean and covariance. However, a stationary background covariance fails to model signal-dependent noise. The adaptive matched subspace detector [10] takes a different approach by modeling background variation with subspaces, but uses a zero mean, white Gaussian signal-independent noise model.

Similar to the noise model implemented by Kerekes and Baum, these algorithms ignore signal-dependent noise under the assumption that its impact is negligible compared to background clutter or that electronics noise is the dominant sensor noise term. However, advances in the design of

photon detectors have minimized the signal-independent electronics noise [11] resulting in sensor noise that is driven by signal-dependent photon noise. Due to these advances it is important to develop an understanding of the impact that this signal-dependent noise has on target detection performance in hyperspectral imagery.

1.2 Research Goals

As part of an effort to model target detection performance of HSI systems, this research assesses the ability to accurately predict the raw signal of a given system as a function of the observed radiance and the optical components comprising the system through the use of a proposed system transmission model. Additionally, a data model is proposed that is intended to accurately portray the effects of signal-dependent sensor noise. This research establishes the framework for an end-to-end system performance model that will be able to predict the target detection performance of a user-defined HSI system for a given target/background scenario, allowing for a high degree of flexibility.

The data used for this research was collected 25-26 June 2014 at Wright Patterson Air Force Base in Dayton, OH. The scene consists of a series of painted wooden panels surrounded by a background of natural foliage alongside other various man-made materials. Noise levels vary between data cubes due to changes in exposure time and the level of solar illumination due to the time of day. Chapter 2 looks at the specifications of the spectrometer, outlines the spectral and radiometric calibration processes, and details the data collection process. Chapter 3 provides insight into the individual elements that make up the HSI system model before defining and investigating the performance of three additive noise models. Chapter 4 provides insight into target detection theory, how it is applied to this research, and derives the detection statistics used to assess the system performance. Chapter 5 steps through many of the simulation and experimental results pertaining to

system modeling and target detection performance. Finally, Chapter 6 outlines the conclusions of the results while putting them into context and opening the door for future research applications.

CHAPTER II

SYSTEM CHARACTERIZATION AND DATA COLLECTION

This chapter highlights the numerous specifications of the hyperspectral system while also providing an overview of the spectral and radiometric calibration processes and results. The chapter concludes with a summary of the experimental set-up including the design/makeup of the scene, the data collection process, and a brief explanation of post-processing procedures.

2.1 Spectrometer Specifications and Implementation

This research uses a Headwall Hyperspec® VNIR E-Series sensor operating in the 400-1000nm range with a quoted spectral resolution of 2-3nm and sampling period of 0.6478$\frac{nm}{pixel}$ [12, 13]. A Schneider Xenoplan 1.9/35mm C-mount lens focuses the light onto the vertically oriented 25µm $\times$ 12mm imaging slit. This lens is ideal due to its high transmission and minimal distortion across the operating range [14]. After transmitting through the lens and slit, light reflects off of a focusing mirror onto a reflective diffraction grating. This separates the light into its spectral components and reflects them on to a focusing mirror and onto the FPA. This FPA is a 2560x2160 array of detectors each with a pitch of 6.5 µm resulting in a 16.64x14.04mm array size. However, due to the physical limitations of the camera, not all of these pixels are used.

The 12mm slit height and unity magnification factor (M_{spec} = 1) result in two inactive sections of the 16.64mm FPA that never receive illumination. These top and bottom sections consist of

355 and 435 inactive rows, respectively. This results in 1770 active spatial rows. However due to collection software limitations, only 1600 spatial rows are collected. Similarly, many of the detectors are ignored in the spectral dimension. While the system is designed for an operating range of 400-1000nm, the FPA experiences incident light from 3-1400nm [13]. Due to limitations

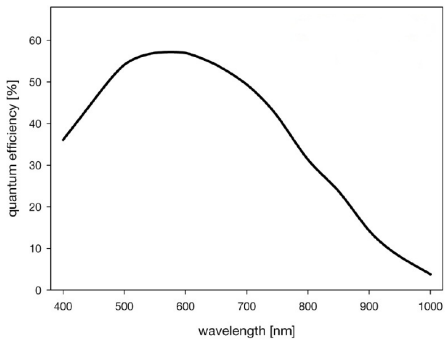


Figure 2.1: Quantum Efficiency of FPA. Headwall Photonics used the above figure to quote the quantum efficiency of the FPA. The spectrometer is designed to operate in the 400-1000nm range. Detectors observing light outside of this range are ignored due to the unknown operating characteristics.

of detector quantum efficiency, grating efficiency, and other optical constraints, the system spectral range is limited to 400-1000nm (Figure 2.1). All detectors with a center wavelength outside of this range are ignored. This also prevents the collection of second order diffraction at some of the longer wavelengths. In order to adequately cover the spectral operating range of the system, 936 spectral bands are collected per frame, each with a sampling period of $0.645\frac{nm}{pixel}$. In summary, the raw data collected from the sensor is 1600 x 936 rather than 2560x2160 before any application specific spectral or spatial trimming is applied.

One limitation of the above design is that the slit removes the horizontal spatial dimension from the field of view (FOV). The second spatial dimension is captured by scanning the spectrometer orthogonal to the slit direction across the desired field of regard. The final result of a hyperspectral

scene collection is called a hyperspectral *datacube* due to its three dimensional nature – two spatial dimensions and one spectral. Figure 2.2 provides a visual representation of a datacube showing the relationship between the spatial and spectral properties. In general, hyperspectral imagery has

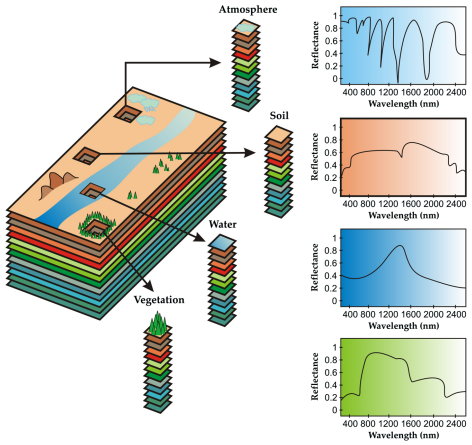


Figure 2.2: Hyperspectral Datacube Visual. The organization and representation of the two spatial dimensions is similar to other forms of imagery. The spectral dimension – represented by the image stacks and plots – contains dense information pertaining to the material makeup of the pixel.

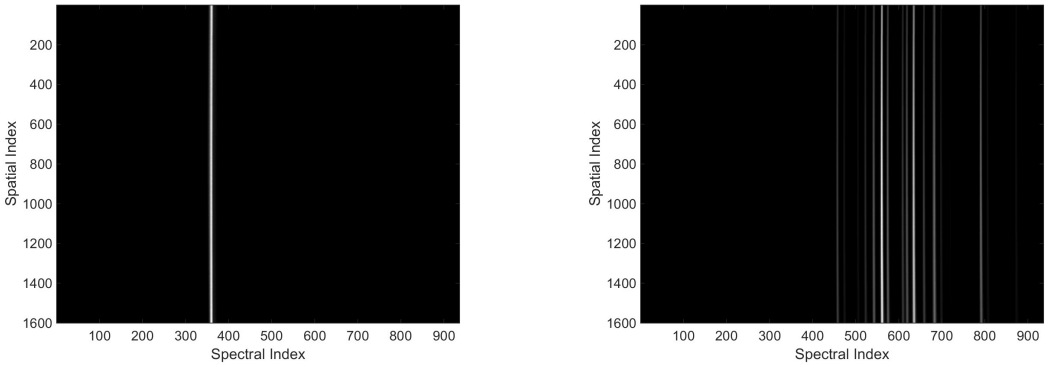
a lower spatial resolution than other remote sensing applications. However, most hyperspectral applications operate using the spectral, rather than spatial, information within the datacube. This means that it is extremely important that the spectral band center represented by each detector is well known.

2.2 Spectral Calibration

Eismann defines spectral calibration as the process of determining the locations of spectral band centers for all samples in a hyperspectral datacube [9]. The locations of these band centers are the result of a function of the optical system in front of the spectrometer's FPA. When an image is recorded each detector corresponds to a spatial location and spectral band. In an ideal system,

spectral information only varies as a linear function of the spectral index k. Due to the numerous optical components necessary for spectrometer design (recall Figure 1.1) it is highly unlikely that an HSI system exhibits perfectly linear characteristics. Spectral calibration is performed in order to achieve an accurate representation of the spectral mapping $\lambda(i, k)$, where λ represents band center.

The spectral calibration process involves collecting datacubes of light sources with well documented spectral characteristics such as lasers and gas lamps. Lasers emit coherent light at a very specific wavelength and narrow bandwidth, while gas lamps have numerous spectral features at various narrow bands with varying intensity. Hyperspectral data from a Helium-Neon (HeNe) laser and an Argon (Ar) gas lamp are compared in Figure 2.3. The image from the HeNe laser shows one



(a) HeNe Laser **(b) Ar Gas Lamp**

Figure 2.3: Spectral Calibration Sources. The left image is an example result from a HeNe laser. One line can clearly be seen and is known to be 632.8 nm. The image on the right is from an Ar gas lamp. Many features of varying intensity can be seen across the FPA.

distinct spectral feature known to be located at 632.8nm [9]. The gas lamp has many more features of varying intensity and bandwidth. The spectral locations as well as the relative strength of these features are documented in the data sheets for the specific gas lamp [15]. The documented wavelength of the features are used along with the observed detector indices to calculate a biquadratic

spectral mapping given by,

$$\lambda(i,k) = ai^2 + bk^2 + cik + di + ek + f, \tag{2.1}$$

where the coefficients a and b describe the level of smile in the spectral and spatial dimensions, c and d account for FPA misalignment, e models linear dispersion, and f describes the spectral offset [9].

The instrument under test, per manufacturer's specifications, is expected to operate in a linear fashion a pixel dispersion of $0.6478\frac{nm}{pixel}$ [13]. Table 2.1 highlights the features used for spectral calibration. These features are chosen to span the operating range of the system. The observed

Table 2.1: Spectral Features of Interest. These features are chosen to span the operating range (400-1000nm) of the imaging system.

Source	Spectral Feature (nm)
HgAr	404.66, 435.84, 578.00
HeNe	623.8
Ar	763.51, 826.45

row-column pairs (i,k) of each feature are recorded while $\lambda(i,k)$ is known for each feature. Using an over determined set of equations, linear regression is used to calculate an estimate of each coefficient. Experimental results indicate that the system under test is described by the coefficient values in Table 2.2.

Table 2.2: Spectral Calibration Coefficients : Biquadratic. Results of spectral calibration assuming a biquadratic model. Near-zero values for a, b, c, and d suggest highly linear performance.

Coefficient	a	b	c	d	e	f
Value	-3.79e-07	2.95e-6	-4.20e-7	7.84e-04	0.6463	399.7893

Referring back to Figure 2.3 shows that these values are not surprising. The images show – from a qualitative perspective – that the results are free from any glaring biquadratic inaccuracies such as smile. This is supported by the low a and b values. Also, there is no noticeable slope, suggesting a well-aligned FPA with low rotation error, which is mathematically supported by the low c and d coefficients. These coefficient values indicate that the spectral dispersion of the system behaves in a highly linear manner.

In an attempt to achieve improved linear fit the process is repeated, but this time a linear mapping is assumed instead of the biquadratic model in (2.1). This simplifies to the following spectral mapping function,

$$\lambda(i,k) = ek + f. \tag{2.2}$$

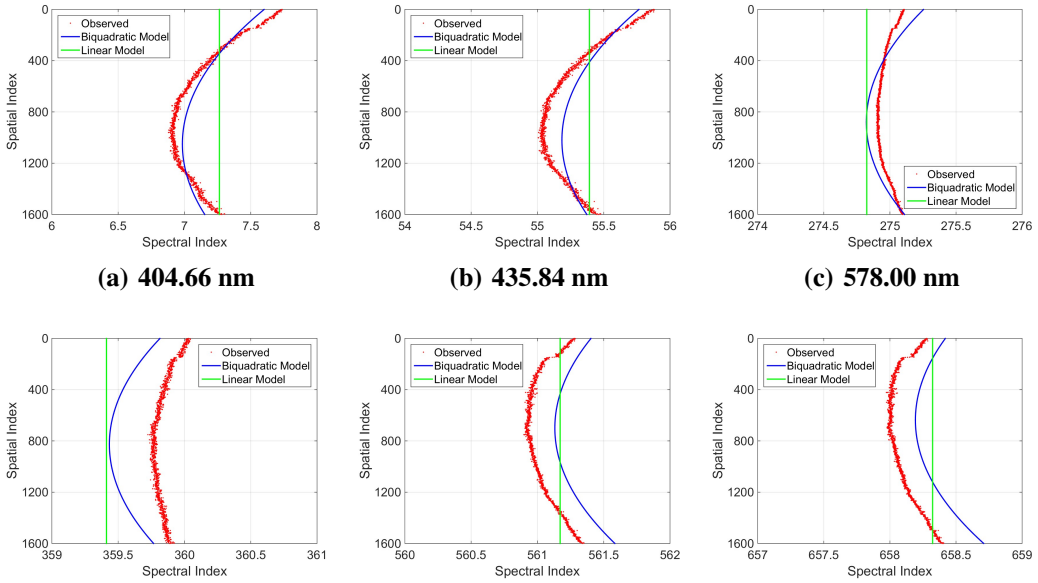
Using the same data and methodology, linear regression is performed resulting in the coefficient values in Table 2.3.

Table 2.3: Spectral Calibration Coefficients : Linear. Results of spectral calibration assuming a linear model. Values are very close to those in Table 2.2 once again hinting at the linearity of the system.

Coefficient	e	f
Value	0.6479	399.95

While the new values of e and f are nearly the same as those calculated in the biquadratic results in Table 2.2, they are a more accurate representation of the linear fit to the data.

Due to the complexity of the biquadratic model, it is reasonable to assume that it would be more thorough and accurate to the true system performance. However, the simplicity of the linear model is attractive. Results from both models are compared in an effort to justify the use of the computationally efficient linear model. Figure 2.4 compares the observed spectral indices of each feature against the predicted indices of both models. Each feature exhibits a small degree of spectral



(a) 404.66 nm **(b) 435.84 nm** **(c) 578.00 nm**

(d) 632.80 nm **(e) 763.51 nm** **(f) 826.45 nm**

Figure 2.4: Spectral Mapping Model Comparison. The red points are the observed spectral band centers from the various features. The blue curve is the feature as described by the biquadratic model while the green line is the same feature according to the linear model. All images are scaled to two spectral indices. At its weakest point (632.8 nm) the linear model is roughly half a pixel off of the observed result.

smile evident from the quadratic nature of the curve in the spectral dimension. Clearly, this is not possible to model with the linear spectral mapping function. The degree of smile varies as a function of wavelength with the strongest observed in the 435.84nm feature. Even in this result, the largest change is slightly less than one pixel over the entire length of the active portion of the FPA. *These results suggest that while the biquadratic model may provide a more accurate approximation to the observed data, the linear model is more than sufficient as it is always within one pixel of the observed value for each of the chosen features.* This linear model is applied to all datasets, accurately mapping a spectral band center to each spectral index of the FPA.

2.3 Radiometric Calibration

Raw image data is the unit-less, digitized result of the observed spectral pupil-plane radiance of the scene after it is subject to a detrimental optical system transfer function. Relative radiometric calibration is often applied to correct for detector response non-uniformity. While this helps to mitigate some of the negative effects of the imaging system, it does not map the raw data to radiometric units. For some scientific applications it is desirable to work in units of spectral radiance – $\frac{\text{Power}}{\text{Unit Area}\times\text{Unit Solid Angle}\times\text{Wavelength}}$ – in order to quantify the observed energy. Absolute radiometric calibration is the process by which raw image data is converted back to radiometric units to represent the amount of energy that entered the HSI system [9].

For this research, a two point absolute radiometric calibration is applied to all data. This process involves the collection of two known reference light levels, declared from here on out as *light* and *dark*. A 2390 ft-Lambert LabSphere® integrating sphere is used as the illumination source for the light reference. The sphere is internally coated with a highly diffuse, highly reflective material and is designed with two access ports: one for an illumination source input and another, larger, exit aperture. The design of the sphere is such that input illumination is reflected multiple times around the interior resulting in a spatially uniform light source at the exit aperture. During the collection of the light frames, the spectrometer is directly aligned with the integrating sphere's exit aperture and a dark cloth is draped over the two instruments to prevent stray light from corrupting the measurements. The output spectral radiance $L(\lambda)$ of the integrating sphere is highly characterized and represented in Figure 2.5 [16]. The dark reference data is collected with a lens cap in place to prevent any incoming light. Hundreds of light and dark frames are recorded and time-averaged in order to minimize effects of noise. These light and dark data are used to calculate a radiometric gain $\hat{a}$ and offset $\hat{b}$ given by,

$$\hat{a}(i,k) = \frac{N_2(i,k) - N_1(i,k)}{L_2\big(\lambda(i,k)\big) - L_1\big(\lambda(i,k)\big)}, \tag{2.3}$$

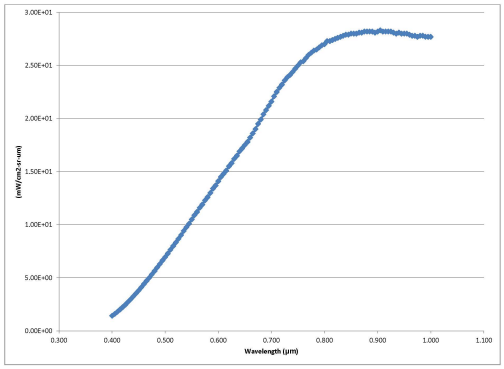


Figure 2.5: Integrating Sphere Output. The system used for this research operates in the 0.4-1.0 µm range. Over this range the integrating sphere displays linearly increasing radiance that levels off before right before 1.0 µm.

and

$$\hat{b}(i,k) = \frac{L_2\big(\lambda(i,k)\big)N_1(i,k) - L_1\big(\lambda(i,k)\big)N_2(i,k)}{L_2\big(\lambda(i,k)\big) - L_1\big(\lambda(i,k)\big)}, \tag{2.4}$$

where N_2 and N_1 are the mean raw light and dark images, and $L_2\big(\lambda(i,k)\big)$, $L_1\big(\lambda(i,k)\big)$ are the known spectral radiance values corresponding to N_2 and N_1, and $\lambda(i,k)$ is defined in (2.2). The use of the lens cap during the collection of the dark frames results in a dark frame reference radiance $L_1\big(\lambda(i,k)\big)$ of zero for all pixels. This simplifies (2.3) and (2.4) to the following two equations,

$$\hat{a}(i,k) = \frac{N_2(i,k) - N_1(i,k)}{L_2\big(\lambda(i,k)\big)}, \tag{2.5}$$

and

$$\hat{b}(i,k) = N_1(i,k). \tag{2.6}$$

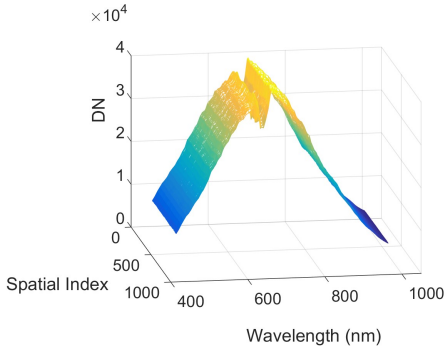
Equation (2.5) is calculated using radiometric data unique to each data collection. The radiometric gain $\hat{a}$ and offset $\hat{b}$ are exploited to calibrate the raw scene data using the following equation,

$$\hat{L}(i,k) = \frac{N(i,k) - \hat{b}(i,k)}{\hat{a}(i,k)}, \tag{2.7}$$

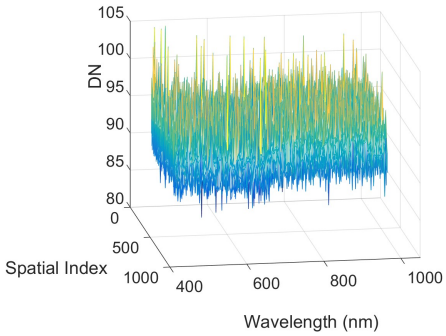
where $N(i,k)$ is the raw scene data and $\hat{L}(i,k)$ is the calibrated data.

Figure 2.6 displays two sample integrating sphere results: one light and one dark. These are the average result of hundreds of frames. Looking back to Figure 2.5 one would expect to observe raw data values in Figure 2.6(a) that increase with wavelength, but the measured results reach a peak at 700nm and steadily decrease at longer wavelengths. This is due to the effects of the detector quantum efficiency seen in Figure 2.1 as well as several other factors that contribute to the optical transfer function. There is an exception to this trend in the spectral range of 670-700nm. In this region, there is a considerable decrease in signal strength. The design of the HSI system is such that there are two separate filters on the FPA, a UV filter and a long pass filter. These two filters have been bonded together and the bond joint is located in this region. This results in a higher absorption factor in the 670-700nm range which is the cause of the sharp decrease in signal strength.

The fact that low signal levels are observed when the true magnitude is known to be high suggests that a high gain will be needed in this region. This is the first of many indicators that this data will be much noisier in the near-infrared than in the visible wavebands. Alternatively, the dark frame data in Figure 2.6(b) does not display any spectral or spatial trends and is narrowly distributed around an average value of 91 Digital Numbers (DN). These results are representative of the 50ms data, but the dark frame results remain consistent with every data cube used for this research – a result consistent with the manufacturer's quoted dark current quote of 2 $\frac{e^-}{sec}$ [17]. The consistent offset of dark noise is likely due to a combination of readout noise and imperfect capacitor draining by the read out circuitry. The absolute radiometric calibration calculations are carried out using the data represented in Figure 2.6. Figure 2.7 shows the results of applying this calibration. In Figure 2.7(a) the radiometric gain $\hat{a}$ is plotted for the entire FPA. The final calibrated result in Figure 2.7(b) is free from the spatial variations observed in the raw data light frame and matches the expected output of the integrating sphere (Figure 2.5).

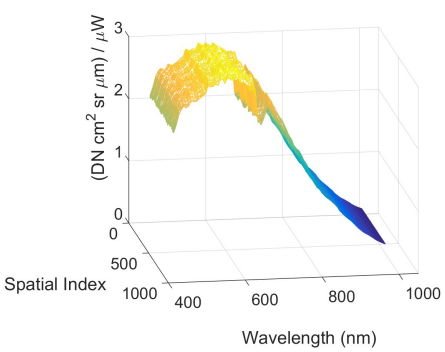


(a) Average Light Frame (N_2)

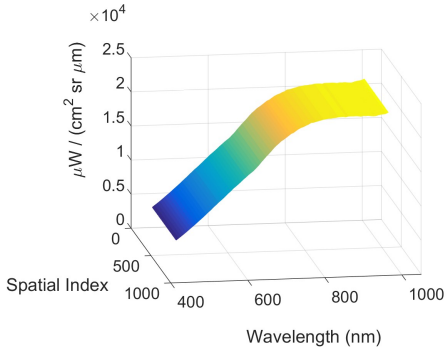


(b) Average Dark Frame (N_1)

Figure 2.6: Average Light and Dark Frames. The light frame displays spectral dependency that is the result of input radiance and system characteristics. Spatial variations due to detector non-uniformity are small in comparison. The absorption effects of the bonding agent can be seen in Figure 2.6(a). The dark frame shows no spatial or spectral dependency.



(a) Radiometric Gain



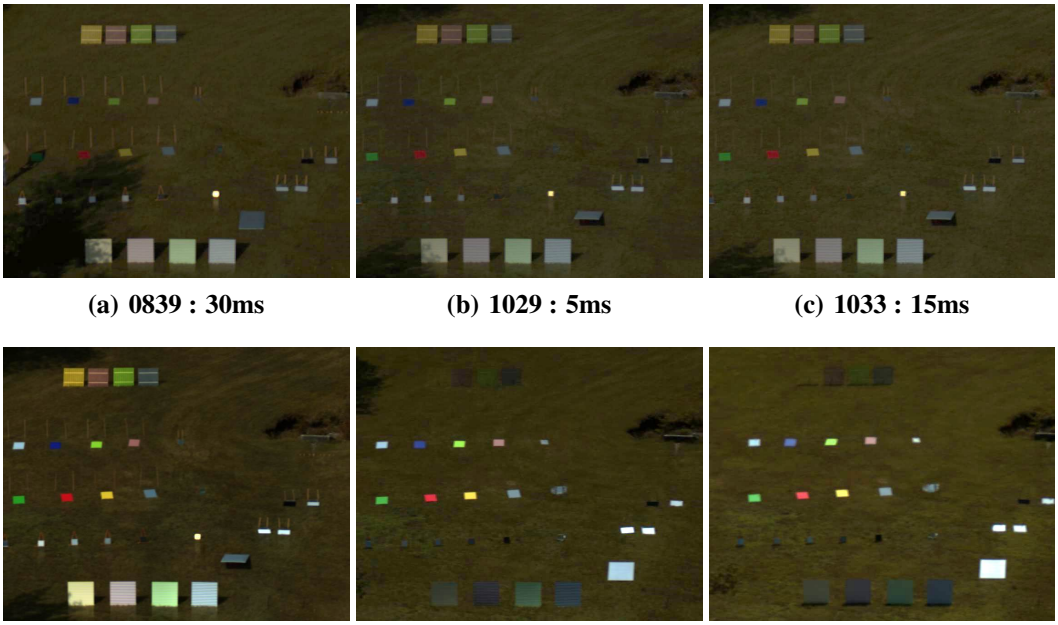
(b) Calibrated Data

Figure 2.7: Radiometric Calibration. Figures show the transformation from raw digital numbers to radiometrically calibrated units. Note that the calibrated data matches the output of the integrating sphere shown in Figure 2.5 and that the effects of the bonding agent have been removed.

2.4 Data Collection Process

The data used for this research was collected during daylight hours of 25-26 June 2014 at Wright Patterson Air Force Base in Dayton, OH. The scene consists of an open, grassy area with several painted panels as viewed from the twelfth floor of a nearby tower. This research investigates six

different data sets – three with varying illumination levels and three with varying exposure times. The three datasets with varying illumination levels are all collected with a 30ms exposure time at three different times of the day: 0839, 1431, and 1624 local time. The 0839 cube was collected on 26 June while the 1431 and 1624 cubes were collected on 25 June. These varying times allow



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure 2.8: True Color Images of Datacubes. Each of these true color images are created from the datacubes used for this research. The change in solar position over time is evident from the tree shadowing in the lower left hand corner of the 0839-1041 images. The targets of interest are the eight painted panels located in the middle, left-hand portion of the scene: light blue, blue, neon green, pink, green, red, yellow, and gray.

for different levels of solar illumination as well as multiple solar angles. The three datasets with varying exposure times were collected in rapid succession in an attempt to observe a constant level of solar illumination. These are collected at 1029, 1033, and 1041 with respective exposure times of 5 ms, 15 ms, and 50 ms. Each of these datacubes are represented in true color form in Figure 2.8. The targets of interest are the eight solid color, painted panels located on the left hand side,

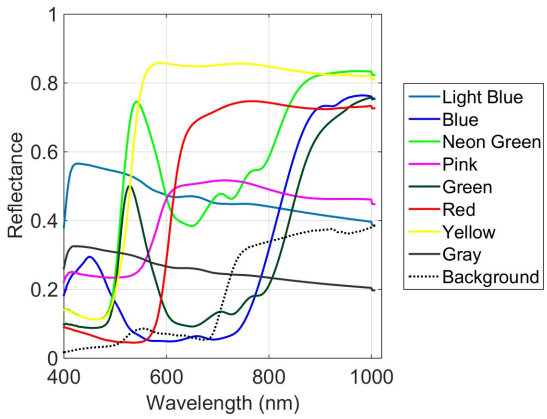


Figure 2.9: Target and Background Reflectance. Spectral reflectance of each target panel. Color of each plot corresponds to the color of the panel in the image of the scene. Average background reflectance is represented by the black dotted line.

vertical middle of the scene organized in two rows of four. Each panel is 2'x4' in size and lies flat on the ground, oriented towards the camera. The spectral reflectance of each of these panels, along with the average reflectance of the background, is recorded in Figure 2.9. Several other man-made materials are present in the scene, however, this research is only concerned with the solid color targets. Therefore, pixels containing these materials are not used for any analysis in this report. A background mask is used to identify grass pixels under uniform solar illumination and free from any of the man-made materials while a target mask for each panel of interest identifies pure target pixels.

Each collection spans a field of regard of roughly 100 x 25 meters. This data was collected in a push-broom manner, panning from left to right using a high precision rotation stage. A measurable amount of time passes from the left side of the datacube to the right. It is desirable during hyperspectral collections to have constant illumination during the recording of a single datacube. When utilizing the push-broom method of data collection, a temporal change in illumination translates to a spatial change in the final data cube which can have a negative impact on target detection and

classification algorithms. To mitigate this issue, it is important to minimize the frame period – the total amount of time needed to record a single frame – so that the possibility of solar illumination variation is minimized as well. Exposure times as low as 1ms and as high as 50ms are used at different stages of the data collection. While the camera possesses flexibility with this parameter, it is unable to operate at a frame rate lower than 30ms. At 1000 frames per collection plus the moving and settling time required by the rotation stage, the total collection time varies between 80 and 100 seconds. The cubes used for this report are free from significant temporal illumination changes.

Immediately after the rotation stage completes its sweep of the scene, radiometric data is collected to be used for absolute radiometric calibration. Collecting this radiometric data immediately following each datacube is meant to achieve an accurate representation of the systems performance at the time of the scene collection, thereby reducing errors due to noise and responsivity drift over time. Before calibrating each datacube, the radiometric data is checked for validity. Some datacubes of integrating sphere data displayed misalignment effects. This occurs when the spectrometer is not aligned perpendicular with the integrating sphere's exit port. Misaligned frames are cut from the radiometric data in order to ensure accuracy. After each cube is calibrated, it is spatially trimmed down to 761x601 spatial pixels. Additionally, due to low SNR, the spectral range is reduced to 455-1000nm with a total of 852 bands. These 852 bands are then down sampled to 57 bands to allow for quicker data modeling and processing.

With the framework of the data collection procedures and the sensor system's spectral and radiometric behavior well understood, the research continues by applying these results to implement a full system transmission model.

CHAPTER III

SYSTEM MODELING IN HYPERSPECTRAL REMOTE SENSING

This chapter looks to outline the theory behind the mathematical model used to calculate this system's transfer function. Experimental tests, simulations, and manufacturer documentation is used to determine several of the system's performance parameters that govern the operating characteristics of this system model. The chapter then introduces signal-dependent noise theory, explains the noise characterization of this system, and defines the signal-dependent noise model used throughout the remainder of the paper.

3.1 System Transmission Modeling

The previous chapter outlined the process of radiometric calibration, the procedure of converting the observed raw digital data into radiometric units through the use of a known light source. An HSI system model addresses the inverse problem. It assumes that the pupil plane radiance is known and looks to predict the raw data result. Whereas radiometric calibration uses an illumination source to calculate its radiometric gain $\hat{a}(i,k)$, the HSI system model uses the physical properties of the optical system. According to Meola et. al. [8] the raw data $N(i,k)$ output by a generic HSI system can be approximated as,

$$N(i,k) = L(i,k)\left(g(i,k)\frac{\pi}{4(f/\#)^2+1}r(i,k)\eta_{QE}(\lambda)\tau(\lambda)\frac{\lambda}{hc}t_d\Delta\lambda A_d\right) + N_1(i,k), \qquad (3.1)$$

where $L(i,k)$ is the pupil plane radiance, $g(i,k)$ is the electron conversion rate, $f/\#$ is the f-number of the system, $r(i,k)$ is the cosine roll-off, $\eta_{QE}(\lambda)$ is the quantum efficiency of the FPA, $\tau(\lambda)$ is the effective transmission of the system, λ is the wavelength of light, h is Planck's constant, c is the speed of light, t_d is the exposure time, $\Delta\lambda$ is the spectral resolution, and A_d is the area of an individual detector. This assumes the bands are narrow enough to approximate, [9]

$$\frac{\pi t_d A_d}{4(f/\#)^2+1}\int L(i,k)\left(g(i,k)r(i,k)\eta_{QE}(\lambda)\tau(\lambda)\frac{\lambda}{hc}\right)d\lambda \approx \ldots$$
$$L(i,k)\left(g(i,k)\frac{\pi}{4(f/\#)^2+1}r(i,k)\eta_{QE}(\lambda)\tau(\lambda)\frac{\lambda}{hc}t_d A_d\right). \quad (3.2)$$

This research attempts to investigate each term of (3.1) in order to calculate a transfer function for this HSI system that will allow for the prediction of raw data from a known incident radiance. Future work would then be able to expand upon this to predict target detection performance as a function of individual system components and operating conditions. The following section investigates each term of this model.

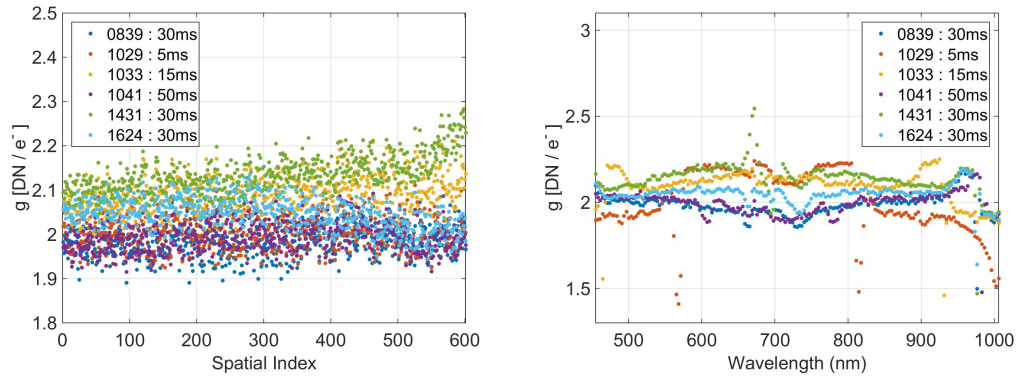
3.1.1 Electron Conversion Estimation

Sperline et. al. implement a practical method of estimating $g(i,k)$ through the use of a mean-variance analysis using raw imaging data [18]. In their work, they plot the observed variance as a function of mean detector response for various light levels. The result is a linear relationship between mean and variance as the light level increases. In Sperline's analysis, the slope of the mean-variance line is shown to estimate the electron conversion rate of the FPA and is given by,

$$\hat{g}(i,k) = \frac{\sigma_2^2(i,k) - \sigma_1^2(i,k)}{\mu_2(i,k) - \mu_1(i,k)}, \quad (3.3)$$

where μ_2 and σ_2^2 are the mean and variance of the detector response at high illumination and μ_1 and σ_2^2 are the statistics resulting from low light data. The electron conversion rate is a direct characteristic of FPA design and this system is expected to operate at a conversion rate of 2.17 $\frac{DN}{e^-}$ [17].

The accuracy of this quote is assessed experimentally using the radiometric calibration data associated with each data collection. Figure 3.1 is a visual representation of the spatial and spectral trends of the electron conversion rate estimated by the application of (3.3). Means are calculated



(a) Spatial Average of g **(b) Spectral Average of g**

Figure 3.1: Estimated Photoelectron Conversion. Figure shows spectral and spatial trends of (3.3) for various datacubes. Figure 3.1(a) plots the average g along each spatial index while Figure 3.1(b) plots the average g along each spectral band. In general, no visual trends are expected however, certain FPA characteristics such as quantum efficiency and A/D conversion accuracy may introduce a degree of nonuniformity in the spectral dimension.

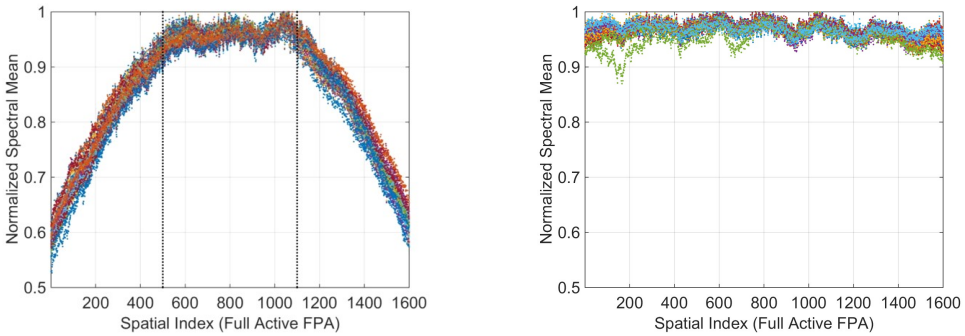
along the spatial and spectral directions in order to find any trends. Four of the six datacubes represented in Figure 3.1(a) do not exhibit any spatial dependency. The remaining two, the 1033 and 1431 datacubes, display a slight increase that is due to misalignment between the spectrometer and the integrating sphere during the collection process rather than truly representing a spatially varying conversion factor. Post-processing techniques aim to remove these misalignment effects, however removing too many frames may have detrimental effects on the radiometric mean and variance calculations.

The spectral means in Figure 3.1(b) suggest that the photoelectron conversion exhibits minimal spectral variation dependent upon the data cube, but the results are centered around the quoted

value of 2.17 $\frac{DN}{e^-}$. It is important to recall Figure 2.6(a). Notice the raw data varies significantly as a function of λ while maintaining near uniformity in the spatial domain. This broad range of values could easily affect the estimates of $g(i,k)$. Over the this range the A/D converters have 'sticking points' where a broad range of analog values are converted to the same digital number. This artificially decreases the variance near these values which in turn decreases the estimated value of g. This effect manifests itself in the local minima seen in each plot. The 0839, 1041, 1431, and 1624 results each display a single minimum near 976nm. The 1029 and 1033 results each show two minima. These are located at 568 and 814nm in the 1029 result and 465 and 931nm in the 1033 result. The cube-to-cube variation is primarily due to varying operating parameters and reference light levels.

3.1.2 Cosine Roll-Off

Lensed remote sensing systems exhibit a spatially varying intensity called vignetting in which the illumination power is greatest at the center of the FPA and falls off in a radial pattern. This illumination roll off is often modeled using a $\cos^4$ function [9] and it represented by $r(i,k)$ in (3.1). For the HSI system under test, this effect only occurs in the spatial direction on the FPA. Spatially uniform integrating sphere data can be used to measure the amount of roll-off observed by the sensor. In this case, any spatial variation observed by the sensor should be due to this phenomena. Figure 3.2 shows raw data from two collections: with and without foreoptics. These results show the effects of cosine roll-off on this system. The roll-off corresponding to the 1041 integrating sphere data is reflective of a f/2.0 lens. The roll-off is symmetric and reduces to 60% at the spatial extremes. Without a lens, cosine roll-off is not present.



(a) Lens On : 1041 Radiometric Data **(b) Lens Off**

Figure 3.2: Cosine Roll-Off Raw Data Comparison. Plots show the effects of cosine roll-off in the system in question. When no lens is used, the response is spatially flat, whereas the intensity reduces to 60% at the spatial extremes with the lens. Note that these plots are representative of the full FPA and that only the data between spatial indices 500 and 1100 are used. This region is highlighted by the vertical lines in Figure 3.2(a). In this region the roll-off is minimal.

3.1.3 Quantum Efficiency and Effective Transmission

The quantum efficiency of a detector represents its probability of converting an incident electron into a detected photon. The performance of this system has been referenced in Figure 2.1. The effective transmission is the term describing all of the cumulative transmission and reflectance terms from the individual elements of the system in question. In the case of this Offner spectrometer the effective transmission can be defined as follows (indexing terms have been removed for simplicity),

$$\tau = (\tau_{optics})(\rho_m^3)(\eta_d)(\tau_{filters}), \tag{3.4}$$

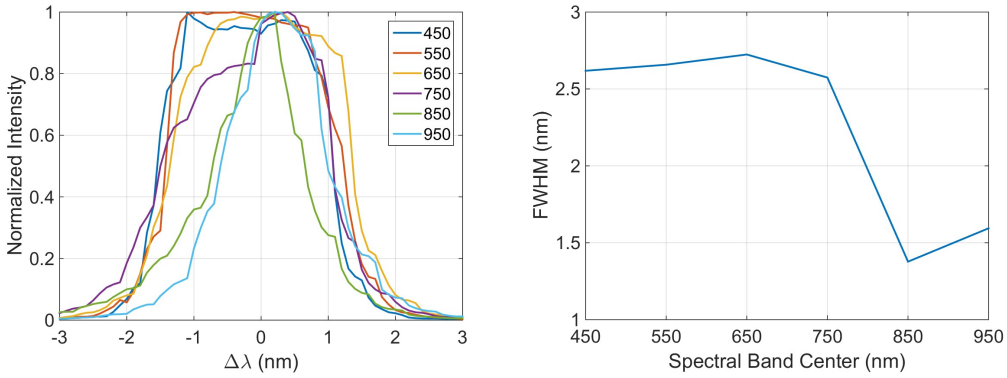
where τ_{optics} is the spectral transmissions of the foreoptic, ρ_m is the reflectance of the mirrors, η_d is the efficiency of the diffraction grating, and $\tau_{filters}$ is a term representing the transmission of all filters in the system. Oftentimes high pass filters are used over portions of the FPA in order to mitigate effects from second order diffraction. Each of these parameters have been provided by the respective manufacturer, but omitted here due to its proprietary nature.

3.1.4 Electron Energy

The term $\frac{\lambda}{hc}$ is the inverted form of the Planck-Einstein coefficient which defines the amount of energy required to form a magnetic field. The linear model outlined in Section 2.2 is used to define λ while hc is constant, leading to a term that increases linearly as a function of wavelength.

3.1.5 Spectral Resolution

The HSI system is quoted as having a spectral resolution between 2-3nm [12]. These results are tested in the lab with a monochromator. Centered around 450nm, 61 measurements are recorded between 453 and 457nm at intervals of 0.1nm. The relative spectral response is recorded in Figure 3.3. This process is repeated at band centers of 550, 650 750, 850, and 950nm. The full width half



(a) Spectral Response **(b) FWHM**

Figure 3.3: Spectral Resolution of HSI System. The spectral response shows similar response shape for 450-750nm while the 850 and 950nm results have narrower characteristics. The FWHM results show that the spectral response behaves as expected until 750nm.

max (FWHM) metric is used to define the spectral resolution and results between 450 and 750nm operate according to the quoted specification while the results outside of this region do not. The low spectral resolution results can be explained by the low detector response at this wavelength.

3.1.6 Constants

All of the terms from (3.1) yet to be defined are constants in the spectral and spatial domains. The f-number is determined by the minimum aperture in the system. The spectrometer is quoted to have an f-number of 2.0 [12]. During the collection, the lens was adjusted to match this. The detector area A_d is also provided by the camera manufacturer. This system has an FPA with a pixel pitch of 6.5 μm [12]. Finally, the exposure time t_d is set by the camera operator at the time of the data collection.

While knowledge of the HSI system's transfer function allows for one to predict camera output, it is based off of an optical model that does not inherently account for variations due to noise. Therefore the result of this model represents the mean data – the result of temporal averaging. The variation due to noise is represented in a separate, signal-dependent additive noise model presented in the following section.

3.2 The Additive Signal-Dependent Noise Model

Consider the observed hyperpixel $\mathbf{x}$, a $K \times 1$ random vector of calibrated spectral radiance values modeled as

$$\mathbf{x} = \mathbf{z} + \mathbf{n}(\mathbf{z}), \tag{3.5}$$

where the hyperpixel $\mathbf{z} = [z_1, z_2, \ldots, z_K]^{\mathrm{T}}$ is the noiseless signal in each of the K sampled spectral bands. Depending upon the underlying target class, $\mathbf{z}$ can be random or deterministic. Sensor noise in each channel is characterized by,

$$\mathbf{n}(\mathbf{z}) = \left[f_1(z_1), f_2(z_2), \ldots, f_K(z_K)\right]^{\mathrm{T}} \tag{3.6}$$

where f_k is a function to be defined later. This definition implies that the sensor noise has both spectrally-dependent and signal-dependent.

The noise characteristics of modern cameras like the one used in this HSI system are thoroughly documented in literature and consist of signal-independent and dependent components that add independently [19]. The signal-independent noise results mainly from electronics noise, a Normally distributed random process. The signal-dependent portion stems from photon shot noise. The photon arrival/detection process is more closely modeled by a Poisson distribution [7]. By definition, a Poisson random variable is characterized by a variance equal to its expected value [20]. It follows that as the mean number of detected photons increases the variance increases linearly. This characteristic allows for the following linear variance model [8],

$$\boldsymbol{\sigma}^2_{\mathbf{n}(\mathbf{z})} = \boldsymbol{\gamma} \odot \mathbf{z} + \boldsymbol{\beta}, \tag{3.7}$$

in which $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are $K \times 1$ vectors of gain and offset values, $\odot$ is the Hadamard product operator, and $\boldsymbol{\sigma}^2_{\mathbf{n}(\mathbf{z})}$ is spectral variance vector of $\mathbf{n}(\mathbf{z})$. The model parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are each a function of the system characteristics described earlier and are unique for every data set. Methods of estimating these parameters from the imagery itself have been explored in literature [11, 21, 22]. However, in an effort to achieve a more robust result, this research utilizes the radiometric calibration data to calculate an empirical estimate of the variance gain and offset at each pixel.

Meola et. al. [8] derive a linear variance model for calibrated hyperspectral imagery utilizing a highly characterized light source that is given by,

$$\sigma^2(i,k) = \frac{g(i,k)}{a(i,k)} L(i,k) + \frac{\sigma_1^2(i,k)}{a^2(i,k)}, \tag{3.8}$$

where $\sigma^2(i,k)$ is the variance of the calibrated data, $g(i,k)$ is the electron conversion rate of the FPA in $\frac{DN}{e^-}$, and $L(i,k)$ is the calibrated data. Section 2.3 defined the estimated radiometric gain $\hat{a}(i,k)$, and $\sigma_1^2(i,k)$ – the dark frame sample variance representing the signal-independent electronics noise– is calculated from the radiometric data.

3.2.1 Noise Parameter Estimation

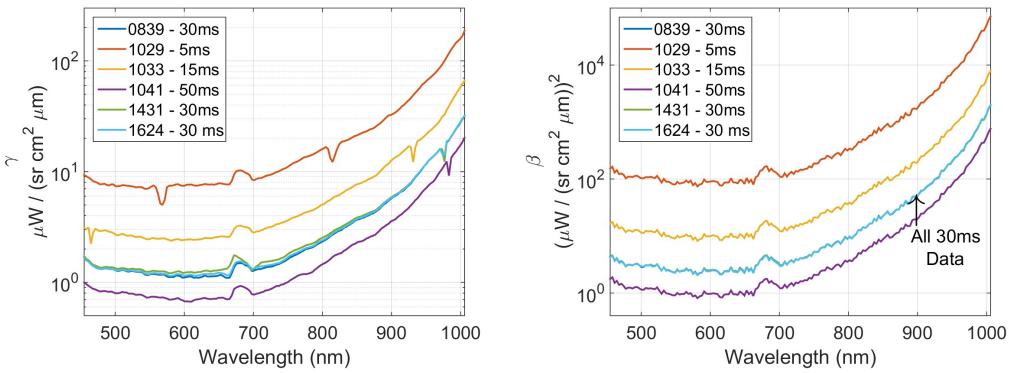
The estimated electron conversion rate $\hat{g}(i,k)$ from (3.3) is substituted into (3.8) to find that the linear variance model parameters $\gamma(i,k)$ and $\beta(i,k)$ are given by,

$$\gamma(i,k) = \frac{g(i,k)}{\hat{a}(i,k)} = \frac{1}{\hat{a}(i,k)} \frac{\sigma_2^2(i,k) - \sigma_1^2(i,k)}{N_2(i,k) - N_1(i,k)}, \tag{3.9}$$

and

$$\beta(i,k) = \frac{\sigma_1^2(i,k)}{\hat{a}^2(i,k)}. \tag{3.10}$$

Noise parameters $\gamma(i,k)$ and $\beta(i,k)$ are calculated according to (3.9) and (3.10). Results indicate that the spatial variation of $\gamma(i,k)$ and $\beta(i,k)$ is minimal for all datacubes used for this research – $\gamma(i,k) \approx \gamma(k)$ and $\beta(i,k) \approx \beta(k)$. Therefore, the spatial index i is ignored and the spectral vectors $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are used for the remainder of the report. Figure 3.4 shows the gain and offset parameters for each of the six datasets used for this research. The noise levels decrease as exposure



(a) Gain ($\boldsymbol{\gamma}$) **(b) Offset ($\boldsymbol{\beta}$)**

Figure 3.4: Spectral Gain and Offset of Additive Noise Model. All results reflect a sharp increase in noise variance towards the longer wavelengths of the operating range. Both parameters are clearly functions of exposure time.

time increases which may initially seem counterintuitive. After all, longer exposure times lead to

higher signal levels which results in increased signal-dependent noise. These results display the role that the radiometric gain plays in the calculation of the noise parameters. Consider the 1029 and 1041 collections. Each observes similar light levels but different exposure times. The cube with the shortest exposure time (1029) has the lowest raw data counts while the 1041 cube (with the longest exposure time) has the highest raw data counts. After calibration, these two cubes should be quite similar due to the similar light levels. This means – according to (2.7) – that the $\hat{a}$ associated with the 1029 cube is much lower than that associated with the 1041 cube. This, when applied to (3.9) and (3.10) and combined with the fact that $g(i,k)$ and $\sigma_1^2(i,k)$ – Figures 3.1 and 2.6(b), respectively – are both similar for all datacubes , explains the noise parameter trends. This also helps explain how the datasets with constant exposure time – 0839, 1431, and 1624 – have nearly identical noise parameters.

In all cases a significant noise increase can be seen in the longer wavelengths. This is due to the low sensor responsivity at these wavelengths. The low responsivity leads to a low $\hat{a}(i,k)$ value in this region (recall Figure 2.7(a)) which results in high variance. In some cases the increasing trend is broken by local irregularities. All of the data sets display a local maxima in the 700nm region due to a bonding agent applied to the top of the FPA. This agent limits the transmission through this spectral region. In addition to this, local minima are seen in the $\boldsymbol{\gamma}$ plots. These features have been passed on from the results of $\hat{g}(i,k)$ in 3.1(b). With the derivations of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$, each parameter of (3.7) is defined and it is possible to derive a model for the distribution of $\mathbf{n}(\mathbf{z})$.

3.2.2 Noise Model Comparison

According to theory, $\mathbf{n}(\mathbf{z})$ is a sum of Poisson and Normal processes and is expected to vary both spatially and spectrally. The work of Foi et. al. [7] is applied with the intention of defining a zero-mean noise term with variance described by (3.7). This process results in the following

distribution of the Poisson + Normal noise model,

$$\mathbf{n}_{PN}(\mathbf{z}) = \mathbf{n}_{\mathcal{P}}(\mathbf{z}) - \mathbf{z} + \mathbf{n}_{SI}, \tag{3.11}$$

where the first term is a Poisson distribution based off of the number of detected electrons converted back into radiometric units and is given by,

$$\mathbf{n}_{\mathcal{P}}(\mathbf{z}) \sim \boldsymbol{\gamma} \odot \mathcal{P}(\boldsymbol{\gamma}^{-1} \odot \mathbf{z}). \tag{3.12}$$

The third term in (3.11) is Normally distributed signal-independent electronics noise given by,

$$\mathbf{n}_{SI} \sim \mathcal{N}(\mathbf{0}, diag(\boldsymbol{\beta})). \tag{3.13}$$

This distribution mixture can make it difficult to accurately model the distribution of $\mathbf{n}_{PN}(\mathbf{z})$. However, as the mean of a Poisson random variable increases its behavior tends towards that of a Normal distribution [7]. In this work an assumption is made that the spectrometer is operating in a photon rich environment. This allows for the shot noise to be approximated by a Normal distribution. This is a commonly used approximation called a *heteroskedastic* noise model and results in the modeled noise behaving according to the following distribution,

$$\mathbf{n}_{H}(\mathbf{z}) \sim \mathcal{N}\big(\mathbf{0}, diag(\boldsymbol{\gamma} \odot \mathbf{z} + \boldsymbol{\beta})\big). \tag{3.14}$$

Here the subscript H denotes the use of the heteroskedastic model. In addition to the Poisson + Normal and heteroskedastic models, the signal-independent noise (3.13) is investigated as it is analogous to many of the currently common noise models in hyperspectral target detection.

A set of simulations are designed to show which of these three noise models is able to most accurately represent the 'true' phenomena by designing three separate detectors based off of the Poisson + Normal (3.11), heteroskedastic (3.14), and signal-independent (3.13) noise models. Each detector is then applied to a set of test data that is distributed according to remote sensing theory. These simulations are designed to cover a realistic scenario in which the noiseless classes (target/background)

exhibit within class variability in addition to sensor noise. Therefore, the two classes used for these simulations are Normally distributed according to,

$$\mathbf{b} \sim \mathcal{N}(\boldsymbol{\mu}_b, \boldsymbol{\Sigma}_b), \tag{3.15}$$

and

$$\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \tag{3.16}$$

where $\mathbf{b}$ and $\mathbf{s}$ are the noiseless background and target and are meant to represent $\mathbf{z}$ in (3.11)-(3.14), $\boldsymbol{\mu}$ is the spectral mean vector and $\boldsymbol{\Sigma}$ is the covariance matrix. it is important to note that this experimental set up is similar, but not identical to the observation model defined later in this research. The definitions in (3.15) and (3.16) are only valid for this investigation into the noise model definitions in this chapter.

Six thousand realizations of $\mathbf{b}$ and $\mathbf{s}$ are generated. Five thousand (N_{test}) of these are selected for the test data set ($\tilde{\mathbf{b}}_{test}$ & $\tilde{\mathbf{s}}_{test}$) and the remaining thousand (N_{train}) are used for training ($\tilde{\mathbf{b}}_{train}$ & $\tilde{\mathbf{s}}_{train}$). Here the tilde denotes a realization of the respective random variable.

Test Data Generation

Test data is generated by using each $\tilde{\mathbf{b}}_{test}$ and $\tilde{\mathbf{s}}_{test}$ to generate a noisy observation $\tilde{\mathbf{x}}_{b,test}$ and $\tilde{\mathbf{x}}_{s,test}$. The application of (3.5) indicates that the distributions of the noisy observations are given by,

$$\mathbf{x}_{b,test} = \tilde{\mathbf{b}}_{test} + \mathbf{n}(\tilde{\mathbf{b}}_{test}), \tag{3.17}$$

and

$$\mathbf{x}_{s,test} = \tilde{\mathbf{s}}_{test} + \mathbf{n}(\tilde{\mathbf{s}}_{test}). \tag{3.18}$$

Recognizing that the Poisson + Normal noise model (3.11) is the accurate noise model according to theory [7], it is applied to (3.17) and (3.18) resulting in the following,

$$\mathbf{x}_{b,test} = \mathbf{n}_{\mathcal{P}}(\tilde{\mathbf{b}}_{test}) + \mathbf{n}_{SI}, \tag{3.19}$$

and

$$\mathbf{x}_{s,test} = \mathbf{n}_{\mathcal{P}}(\tilde{\mathbf{s}}_{test}) + \mathbf{n}_{SI}. \tag{3.20}$$

In practice, each noiseless realization is used to generate a single Poisson realization which is then added to a single realization of a Normal distribution. This results in a single observation sample. The process is repeated for each $\tilde{\mathbf{b}}_{test}$ and $\tilde{\mathbf{s}}_{test}$ resulting in the noisy test data set.

Training Data Generation

The three detectors used on the test data are trained with identical noiseless data sets ($\tilde{\mathbf{b}}_{train}$ & $\tilde{\mathbf{s}}_{train}$) and differ only in their additive noise model: signal-independent, heteroskedastic, and Poisson + Normal. The signal-independent training set adds a Normally distributed signal-independent noise term (3.13) resulting in observation classes given by,

$$\mathbf{x}_{b,SI} = \tilde{\mathbf{b}}_{train} + \mathbf{n}_{SI}, \tag{3.21}$$

and

$$\mathbf{x}_{s,SI} = \tilde{\mathbf{s}}_{train} + \mathbf{n}_{SI}. \tag{3.22}$$

The heteroskedastic model adds signal-dependent, zero-mean, Normally distributed noise as seen in (3.14) to the noiseless realizations resulting in the distributions given by,

$$\mathbf{x}_{b,H} = \tilde{\mathbf{b}}_{train} + \mathbf{n}_{H}(\tilde{\mathbf{b}}_{train}), \tag{3.23}$$

and

$$\mathbf{x}_{s,H} = \tilde{\mathbf{s}}_{train} + \mathbf{n}_{H}(\tilde{\mathbf{s}}_{train}). \tag{3.24}$$

In practice, the observations defined in (3.21) - (3.24) are created by adding a single Normally distributed realization to each noiseless realization. Finally, in a result analogous to (3.19) and (3.20), the Poisson + Normal model is distributed,

$$\mathbf{x}_{b,PN} = \mathbf{n}_{\mathcal{P}}(\tilde{\mathbf{b}}_{train}) + \mathbf{n}_{SI}, \tag{3.25}$$

and

$$\mathbf{x}_{s,PN} = \mathbf{n}_{\mathcal{P}}(\tilde{\mathbf{s}}_{train}) + \mathbf{n}_{SI}. \tag{3.26}$$

This results in 1000 background and 1000 target noisy observations for each of the three training sets. After the training sets are created, two non-parametric detection algorithms – K-Nearest Neighbors (Knn) [23] and Parzen [24] – are applied to each. These algorithms attempt to fit a distribution to the training data. The trained result from each of the three additive noise models is then used as a detection statistic on the test data. A Receiver Operating Characteristic (ROC) curve [25] is generated for each result to compare pD as a function of FAR.

Test Design

Due to the complexity of the models in (3.21) - (3.26), there are a high number of parameters to adjust. The goal is to measure performance trends as a function of signal mean ($\boldsymbol{\mu}_b$ & $\boldsymbol{\mu}_s$). This is due to the fact that the Poisson + Normal model is expected to tend towards the Heteroskedastic model as mean signal increases. Additionally, the signal-dependent portion of the noise increases with mean signal. This implies that the accuracy of the signal-independent model (3.21) & (3.22) will decrease as signal mean increases. Multiple assumptions and simplifications are made in order to isolate $\boldsymbol{\mu}_b$ as the lone variable. To start, each observation has only two bands. The low dimensionality leads to simplified calculations while also allowing the data to be visualized with scatter plots (Figure 3.5). For both classes the mean of the first band is equal to the mean of the second band,

$$\boldsymbol{\mu}_b(1) = \boldsymbol{\mu}_b(2), \tag{3.27}$$

and

$$\boldsymbol{\mu}_s(1) = \boldsymbol{\mu}_s(2). \tag{3.28}$$

Additionally, features for both classes are independently distributed, given by a diagonal covariance matrix ($\boldsymbol{\Sigma}_b$ & $\boldsymbol{\Sigma}_s$). Finally, class separation must remain constant due to its role in detector performance. A constant separability measure d_B is defined and the Bhattacharyya distance [26],

$$d_B = \frac{1}{8}(\boldsymbol{\mu}_b - \boldsymbol{\mu}_s)^{\mathrm{T}} \left(\frac{\boldsymbol{\Sigma}_b + \boldsymbol{\Sigma}_s}{2} \right)^{-1} (\boldsymbol{\mu}_b - \boldsymbol{\mu}_s) + \frac{1}{2} \ln \left(\frac{\left| \frac{\boldsymbol{\Sigma}_b + \boldsymbol{\Sigma}_s}{2} \right|}{\sqrt{|\boldsymbol{\Sigma}_b||\boldsymbol{\Sigma}_s|}} \right), \tag{3.29}$$

is used to calculate the target class mean $\boldsymbol{\mu}_s$ that satisfies (3.29). It's important to note that this is the separability of the noiseless class distributions and therefore the added noise will reduce the separability between the classes. Table 3.1 shows the settings used for each of the simulations, where N_{test} is the number of test realizations per iteration, N_{train} is the number of training realizations per iteration, and N_m is the number of iterations averaged together in Monte Carlo fashion in order to arrive at the final results. Three values of $\boldsymbol{\gamma}$ – $[1,1]^{\mathrm{T}}$, $[2,2]^{\mathrm{T}}$, & $[3,3]^{\mathrm{T}}$ – are used to influence the

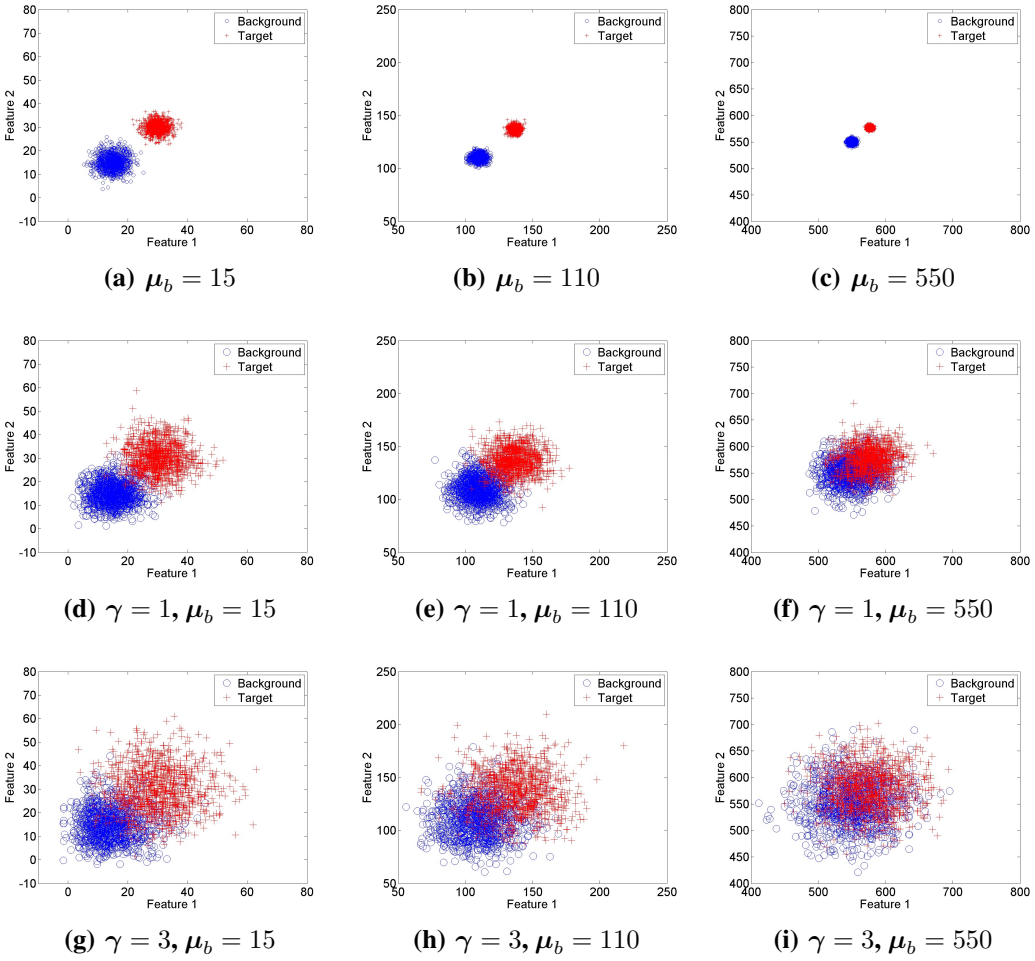
Table 3.1: Simulation Parameters Used for Noise Model Comparison. These are the parameters that remain constant for all noise model comparison simulations.

Variable	$\boldsymbol{\Sigma}_b$	$\boldsymbol{\Sigma}_s$	D_B	N_{test}	N_{train}	N_m	$\boldsymbol{\beta}$
Value	$\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	60	5000	1000	30	$\begin{bmatrix} 5 \\ 5 \end{bmatrix}$

level of signal-dependent noise between simulations.

Results

Figure 3.5 shows the scatter plots for several sets of training data simulations. The top row of scatter plots represent $\mathbf{x}_{b,SI}$ and $\mathbf{x}_{s,SI}$ for varying levels of $\boldsymbol{\mu}_b$ while the next two rows represent $\mathbf{x}_{b,PN}$ and $\mathbf{x}_{s,PN}$ for various $\boldsymbol{\mu}_b$ and $\boldsymbol{\gamma}$ values. The heteroskedastic model is not represented due to its visual similarity to the Poisson + Normal represented in the bottom two rows. The first column shows results for the lowest signal level used, $\boldsymbol{\mu}_b = 15$. From a visual perspective, all three examples in this column have two separable classes,but this separability decreases as the noise gain



(a) $\mu_b = 15$ **(b)** $\mu_b = 110$ **(c)** $\mu_b = 550$

(d) $\gamma = 1$, $\mu_b = 15$ **(e)** $\gamma = 1$, $\mu_b = 110$ **(f)** $\gamma = 1$, $\mu_b = 550$

(g) $\gamma = 3$, $\mu_b = 15$ **(h)** $\gamma = 3$, $\mu_b = 110$ **(i)** $\gamma = 3$, $\mu_b = 550$

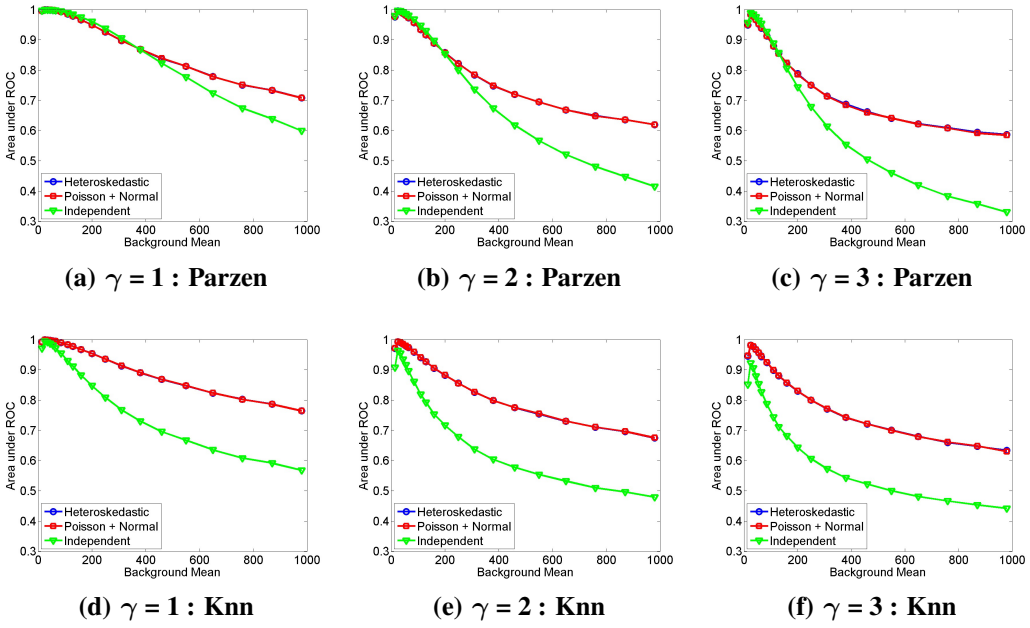
Figure 3.5: Noise Model Scatter Plots. This figure shows the scatter plots for several sets of training data. The top row represents samples from the signal-independent noise model while the next two rows represent the Poisson + Normal signal-dependent model. The signal-dependent trends suggest that the two classes become harder to separate as the signal level (μ_b) increases. The heteroskedastic model is not represented due to its visual similarity to the Poisson + Normal results.

γ increases. As μ_b increases, the disparity between the signal-independent and signal-dependent models grows greater. This is especially evident when comparing Figures 3.5(c) and 3.5(i). Two easily separable distributions in Figure 3.5(c) become difficult to distinguish in the presence of signal-dependent noise (Figure 3.5(i)).

The goal of this experiment is to show how closely each of the three noise models is able to model the theoretical noise characteristics. It is expected that the detector trained with the Poisson + Normal noise data would outperform the other two models since it is distributed identically to the test data. The heteroskedastic model detector is expected to perform slightly worse than the Poisson + Normal at low signal levels, while closing this performance gap as the signal means increase. Finally, the signal-independent result is expected to perform poorly, especially at high signal levels and when high levels of noise gain γ are used. While these results are representative of training data, it is necessary to point out that Figures 3.5(d)-3.5(i) are identically distributed with the corresponding test data. Keeping this in mind, it is clear that the signal-independent model performance will suffer as γ increases. For instance, a decision statistic based off of Figure 3.5(a) will perform better on the data in Figure 3.5(d) than that in Figure 3.5(g). As the mean increases, the performance is expected to drop as well. The decision statistic formed from the data in Figure 3.5(c) clearly is not an optimal result for Figures 3.5(f) and 3.5(i).

Figure 3.6 looks at the results of the simulations for a range of signal levels. The performance metric used is the area under ROC, which is a method of representing a ROC curve's performance as a single scalar value. An area under ROC value of 1 indicates that for all possible τ the detector reports 100% detection and zero false alarms. Lower values indicate decreased performance.

According to the Parzen classifier the signal-independent model is the best performer for low signal applications. As γ increases the signal-dependent models outperform the signal-independent in an increasingly significant manner. This is due to the increased influence of signal-dependent noise. On the other hand, the Knn results suggest that the signal-independent model never outperforms the other two models. Together, these results suggest that there may be some low signal cases in which a signal-independent model may be optimal. However, both algorithms agree that



(a) $\gamma = 1$: Parzen **(b) $\gamma = 2$: Parzen** **(c) $\gamma = 3$: Parzen**

(d) $\gamma = 1$: Knn **(e) $\gamma = 2$: Knn** **(f) $\gamma = 3$: Knn**

Figure 3.6: Noise Model Performance Results. The top row shows results from the Parzen algorithm while the bottom represents Knn results. All examples exhibit decreasing performance as the signal mean increases. This is due to decreasing separation between the noisy classes evident in Figure 3.5. It can be seen that both of the signal-dependent noise models perform nearly identically for all signal levels and in most cases perform better than the signal-independent model.

in photon-rich environments signal-dependent models can be expected to outperform their signal-independent counterparts. *The results also suggest that the Heteroskedastic model is an accurate approximation of the Poisson + Normal model for all tested signal levels*. While these simulations are limited in scope due to the number of assumptions made, these results justify the use of the Heteroskedastic noise model throughout the remainder of the report.

This chapter has established the framework of the revised sensor system and noise models. The following chapter applies this noise model to derive a signal dependent detection statistic that will be used to determine the target detection performance of the system throughout the remainder of the research.

CHAPTER IV

TARGET DETECTION THEORY AND APPLICATION

This chapter outlines the basics of detection theory by defining likelihood ratio tests and decision thresholding. The observation model is defined and used to calculate a signal-dependent detection statistic. A conditional probability density function of this detection statistic is also defined before the derivation of its first and second order statistics.

4.1 Detection Theory

One of the many applications of hyperspectral remote sensing is target detection, the detection of a specific material within a scene. Target detection can be thought of as a two class classification problem between target and background. Due to the possibility of natural variability within either class, it is usual to assume that one or more of these classes behave according to a known statistical distribution. A common assumption – and one that is used throughout this paper – is that a single class can be accurately modeled by the multivariate Normal model given by,

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{K}{2}}} \frac{1}{|\mathbf{\Sigma_x}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_\mathbf{x})^{\mathrm{T}}\mathbf{\Sigma}_\mathbf{x}^{-1}(\mathbf{x}-\boldsymbol{\mu}_\mathbf{x})}, \tag{4.1}$$

where $\mathbf{x}$ is the observed, noisy hyperpixel defined in Section 3.2, $\boldsymbol{\mu}_\mathbf{x}$ and $\mathbf{\Sigma_x}$ are the mean and covariance of $\mathbf{x}$, and $p(\mathbf{x})$ is the probability of $\mathbf{x}$ occurring.

The target detection process consists of testing each pixel to determine whether it is more likely to belong to the target or background class. This begins with the application of the likelihood ratio test (LRT) given by,

$$l(\mathbf{x}) = \frac{p(\mathbf{x}|H_1)}{p(\mathbf{x}|H_0)}, \tag{4.2}$$

where $p(\mathbf{x}|H_1)$ is the conditional probability density function (PDF) given that $\mathbf{x}$ is of the target class and $p(\mathbf{x}|H_0)$ is the conditional PDF given that $\mathbf{x}$ is of the background class. As $l(\mathbf{x})$ increases the likelihood that $\mathbf{x}$ belongs to the target class increases as well. When both conditional PDFs are Normally distributed, $l(\mathbf{x})$ is rewritten,

$$l(\mathbf{x}) = \frac{|\boldsymbol{\Sigma}_{\mathbf{x}_0}|^{\frac{1}{2}}}{|\boldsymbol{\Sigma}_{\mathbf{x}_1}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_1})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{x}_1}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_1})+\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_0})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{x}_0}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_0})}, \tag{4.3}$$

where the subscript 0 or 1 indicate which hypothesis the metric falls under. This is further simplified by taking the logarithm of both sides and ignoring constant terms,

$$\ln\big(l(\mathbf{x})\big) = (\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_0})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{x}_0}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_0}) - (\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_1})^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{x}_1}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{x}_1}). \tag{4.4}$$

The desired final result of a detection algorithm is a binary image in which each pixel has been declared to be either target or background. Therefore, a threshold must be applied to (4.4) in the following manner,

$$\theta(\mathbf{x}) = \begin{cases} 1, & \ln\big(l(\mathbf{x})\big) \geq \tau \\ 0, & \ln\big(l(\mathbf{x})\big) < \tau \end{cases} \tag{4.5}$$

where τ is the threshold designed to achieve the desired performance and a value of 1 indicates the classification of target while 0 is the classification for background hyperpixels. The method of calculating τ is as important as the design of the statistical class models when it comes to target detection algorithm design. Changing the value of τ has a direct effect on the pD and FAR. The 'optimal' value of τ is often application dependent. Many hyperspectral applications design a τ

to result in a desired constant false alarm rate (CFAR). This method is modeled after the Neyman-Pearson hypothesis testing method [27] and is given by,

$$\int_{\tau}^{\infty} p(l(\mathbf{x})|H_0) = \zeta, \tag{4.6}$$

where ζ is the desired CFAR. This method requires prior knowledge of the conditional distribution of $l(\mathbf{x})$ under the null hypothesis in order to be applicable. Detection statistic results throughout this report are thresholded using this method.

4.2 Observation Model

Consider a hypothesis testing problem for full-pixel targets given by :

$$\begin{array}{ll} H_0: & \mathbf{x} = \mathbf{b} + \mathbf{n}_0 \\ H_1: & \mathbf{x} = \mathbf{s} + \mathbf{n}_1 \end{array} \tag{4.7}$$

where $\mathbf{b}$ is the background, $\mathbf{s}$ is the target, and $\mathbf{n}_0$ and $\mathbf{n}_1$ are sensor noise under each respective hypothesis. Due to natural variations in the background class (in this case, grass), $\mathbf{b}$ is assumed to be Normally distributed with mean $\boldsymbol{\mu}_b$ and covariance $\boldsymbol{\Sigma}_b$, both assumed to be known. It is important to note that in real data $\mathbf{b}$ will likely *not* be Normally distributed and this assumption is made for simplicity. The targets are freshly painted panels expected to have little to no spatial/spectral variability. Therefore, $\mathbf{s}$ is assumed to be deterministic and known. The noise terms ($\mathbf{n}_0$ and $\mathbf{n}_1$) are modeled with the heteroskedastic model defined in (3.14), where the generic noiseless hyperpixel $\mathbf{z}$ is a realization of the background $\mathbf{b}$ for $\mathbf{n}_0$ or the deterministic target spectra $\mathbf{s}$ for $\mathbf{n}_1$. As the sum of two Normally distributed, independent random vectors, $\mathbf{x}$ is also Normally distributed under each hypothesis.

$$\begin{array}{ll} H_0: & \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_b, \boldsymbol{\Sigma}_0) \\ H_1: & \mathbf{x} \sim \mathcal{N}(\mathbf{s}, \boldsymbol{\Sigma}_1) \end{array} \tag{4.8}$$

where the covariance under the null hypothesis is defined,

$$\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}_b + \boldsymbol{\Sigma}_{n_0} = \boldsymbol{\Sigma}_b + diag\big(\boldsymbol{\gamma} \odot \tilde{\mathbf{b}} + \boldsymbol{\beta}\big), \tag{4.9}$$

where $\tilde{\mathbf{b}}$ denotes a realization of the random variable $\mathbf{b}$. It follows that the covariance under the alternate hypothesis is given by,

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_{n_1} = diag(\boldsymbol{\gamma} \odot \mathbf{s} + \boldsymbol{\beta}). \tag{4.10}$$

By applying (4.7), (4.9), and (4.10) to (4.4), the log LRT resulting from this observation model is found to be,

$$\ln\big(l(\mathbf{x})\big) = (\mathbf{x} - \boldsymbol{\mu}_b)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) - (\mathbf{x} - \mathbf{s})^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}). \tag{4.11}$$

For notational purposes, $\ln\big(l(\mathbf{x})\big)$ is simplified to $l(\mathbf{x})$ and from this point forward $l(\mathbf{x})$ represents an unthresholded log LRT. The following section outlines how each term of (4.11) is calculated from the real data and derives two of the detection statistics implemented throughout the remainder of the report.

4.3 Model Application

Recall from (4.9) that $\boldsymbol{\Sigma}_0$ is a function of a noiseless background realization $\tilde{\mathbf{b}}$. In many real-world applications $\tilde{\mathbf{b}}$ is not easily found from the noisy observation $\mathbf{x}$. In this work $\tilde{\mathbf{b}}$ is approximated by the noisy observation $\mathbf{x}$ for real data analysis. While this may be a good estimate while $\mathbf{x}$ is of H_0, it is a very poor estimate when $\mathbf{x}$ is of H_1. According to (4.7), this estimate is also dependent upon the signal-to-noise ratio (SNR) of the data. For the high SNR data used for this research it is assumed to be a good estimate, but it may prove problematic for low SNR applications.

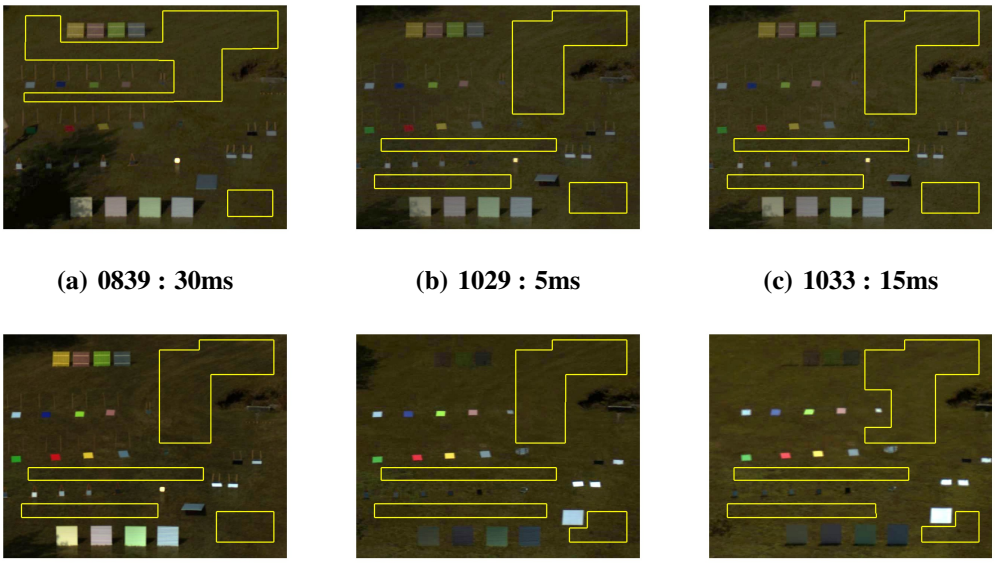
This research is primarily concerned with the behavior of $p(l(\mathbf{x})|H_0)$ and therefore it is known *a priori* that each $\mathbf{x}$ is of H_0, making $\tilde{\mathbf{b}} \approx \mathbf{x}$ a valid approximation. However, the inherent nature of real world target detection applications is such that the class of $\mathbf{x}$ is not known *a priori*. In these applications it may be necessary to perform some sort of preprocessing with an anomaly detector such as the Reed-Xiaoli (RX) algorithm [28]. A pixel with a low RX value is spectrally similar to the rest of the image and is therefore assumed to be similar to the background. This information

could be used to pre-sort the scene and establish some decision rule such as,

$$\tilde{\mathbf{b}} \approx \begin{cases} \bar{\mathbf{b}}, & RX(\mathbf{x}) \geq \phi \\ \mathbf{x}, & RX(\mathbf{x}) < \phi \end{cases} \tag{4.12}$$

and for all anomalous $\mathbf{x}$ as determined by an arbitrary threshold, $\tilde{\mathbf{b}}$ could then be approximated by $\bar{\mathbf{b}}$. This preprocessing step is presented as a proposal for future real world application as, once again, this research focuses on the behavior of the detection statistics under the null hypothesis only.

Another real-world consequence is that $\boldsymbol{\mu}_b$ and $\boldsymbol{\Sigma}_b$ are not truly known and must be estimated from the data itself. Background masks (Figure 4.1) are used to identify the scene's background pixels and calculate the sample mean $\hat{\boldsymbol{\mu}}_b$ and covariance $\hat{\boldsymbol{\Sigma}}_0$. Note that the sample covariance matrix



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure 4.1: Background Contours of Tested Images. The yellow contour in each image surrounds the background truth pixels. For five of the six, samples are collected from the same regions. The tree shadow located in the lower left corner of 0839 prevent those grass pixels from being used. In each case, the background class consists of 93,204 pixels.

$\hat{\boldsymbol{\Sigma}}_0$ estimated from real background data contains both the background covariance and sensor noise.

Here the assumption is made that the sensor noise in $\hat{\boldsymbol{\Sigma}}_0$ is driven by the mean background signal $\hat{\boldsymbol{\mu}}_b$. Using (4.9) the estimated background covariance is calculated,

$$\hat{\boldsymbol{\Sigma}}_b = \hat{\boldsymbol{\Sigma}}_0 - \bar{\boldsymbol{\Sigma}}_{n_0}, \tag{4.13}$$

where $\bar{\boldsymbol{\Sigma}}_{n_0}$ is defined,

$$\bar{\boldsymbol{\Sigma}}_{n_0} = diag\big(\boldsymbol{\gamma} \odot \hat{\boldsymbol{\mu}}_b + \boldsymbol{\beta}\big). \tag{4.14}$$

The target signal $\mathbf{s}$ is the sample mean of the target under test calculated from the data using a separate mask. Mask sizes used to calculate $\mathbf{s}$ for each target vary in size from 135 to 228 pixels, which should provide an adequate noiseless estimate.

In practice, the theory-based detection statistic from (4.11) is applied as,

$$l_{dep}(\mathbf{x}) = (\mathbf{x} - \hat{\boldsymbol{\mu}}_b)^{\mathrm{T}}\big(\hat{\boldsymbol{\Sigma}}_b + diag(\boldsymbol{\gamma} \odot \hat{\hat{\mathbf{b}}}(\mathbf{x}) + \boldsymbol{\beta})\big)^{-1}(\mathbf{x} - \hat{\boldsymbol{\mu}}_b) - (\mathbf{x} - \mathbf{s})^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}(\mathbf{x} - \mathbf{s}), \tag{4.15}$$

where $\hat{\hat{\mathbf{b}}}(\mathbf{x})$ is an estimate of the background realization – for this work $\hat{\hat{\mathbf{b}}}(\mathbf{x}) = \mathbf{x}$. For performance comparison a signal-independent detection statistic is also defined. In this model, the additive noise term is modeled by the mean noise variance $\bar{\boldsymbol{\Sigma}}_{n_0}$ rather than the pixel-dependent diagonal matrix from (4.15). This simplifies the result to the following,

$$l_{ind}(\mathbf{x}) = (\mathbf{x} - \hat{\boldsymbol{\mu}}_b)^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_0^{-1}(\mathbf{x} - \hat{\boldsymbol{\mu}}_b) - (\mathbf{x} - \mathbf{s})^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}(\mathbf{x} - \mathbf{s}). \tag{4.16}$$

4.4 Theoretical Distributions

The result from (4.6) shows that an accurate model of $l(\mathbf{x})|H_0$ is needed in order to properly design a τ to result in a desired CFAR. In general, the conditional PDF of $l(\mathbf{x})$ under either hypothesis will most likely be non-Normal. However, the central limit theorem may be applicable in certain cases with high enough dimensionality to produce a result that may be approximated by a Normal distribution [26]. This work assumes that the central limit theorem applies to the results due

to the high dimensionality of hyperspectral data and makes note of instances where the assumption appears invalid.

By assuming a Normal distribution, only the first and second order statistics are needed to fully model the conditional PDF. The mean and variance of (4.11) are derived for this purpose. Before the calculations are carried out it is helpful to present four properties of expectations of random variables. These identities are used extensively throughout the derivations.

Definition If $\mathbf{x}$ is a $K \times 1$ random vector with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, $\mathcal{A}$ is a deterministic matrix, and $\mathbf{a}$, $\mathbf{b}$ are deterministic $K \times 1$ vectors then the following are true [29, 30].

$$E\big[\mathbf{x}^{\mathrm{T}}\mathcal{A}\mathbf{x}\big] = Tr(\mathcal{A}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^{\mathrm{T}}\mathcal{A}\boldsymbol{\mu} \tag{4.17}$$

$$E\big[(\mathbf{a}^{\mathrm{T}}\mathbf{x})(\mathbf{b}^{\mathrm{T}}\mathbf{x})\big] = \mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{b} + (\mathbf{a}^{\mathrm{T}}\boldsymbol{\mu})(\mathbf{b}^{\mathrm{T}}\boldsymbol{\mu}) \tag{4.18}$$

$$E\big[(\mathbf{x}^{\mathrm{T}}\mathcal{A}\mathbf{x})(\mathbf{a}^{\mathrm{T}}\mathbf{x})\big] = 2\boldsymbol{\mu}^{\mathrm{T}}\mathcal{A}\boldsymbol{\Sigma}\mathbf{a} + Tr(\mathcal{A}\boldsymbol{\Sigma})(\mathbf{a}^{\mathrm{T}}\boldsymbol{\mu}) + (\mathbf{a}^{\mathrm{T}}\boldsymbol{\mu})(\boldsymbol{\mu}^{\mathrm{T}}\mathcal{A}\boldsymbol{\mu}) \tag{4.19}$$

$$\begin{aligned} &E\big[(\mathbf{x}^{\mathrm{T}}\mathcal{A}_1\mathbf{x})(\mathbf{x}^{\mathrm{T}}\mathcal{A}_2\mathbf{x})\big] = \\ &2Tr(\mathcal{A}_1\boldsymbol{\Sigma}\mathcal{A}_2\boldsymbol{\Sigma}) + 4(\boldsymbol{\mu}^{\mathrm{T}}\mathcal{A}_1\boldsymbol{\Sigma}\mathcal{A}_2\boldsymbol{\mu}) + \big(\boldsymbol{\mu}^{\mathrm{T}}\mathcal{A}_1\boldsymbol{\mu} + Tr(\mathcal{A}_1\boldsymbol{\Sigma})\big)\big(\boldsymbol{\mu}^{\mathrm{T}}\mathcal{A}_2\boldsymbol{\mu} + Tr(\mathcal{A}_2\boldsymbol{\Sigma})\big) \end{aligned} \tag{4.20}$$

In (4.17)-(4.20) $Tr()$ is the matrix trace operator. It is important to note here that these definitions assume deterministic matrices. The nature of $\boldsymbol{\Sigma}_0$ is that of a random matrix dependent upon the observed background realization. However, these calculations are implemented after the observation of $\mathbf{x}$ at which point $\boldsymbol{\Sigma}_0$ is no longer random. Therefore, these equations may be accurately applied. The mean and variance of the expected distribution must be calculated separately for each

hypothesis since the characteristics of $\mathbf{x}$ change accordingly. The derivation begins by assuming H_0 and rewriting (4.11) as follows,

$$l(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu}_b)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) - \mathbf{x}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{x} + 2(\mathbf{s}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{x}) - \mathbf{s}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{s}. \tag{4.21}$$

Under the null hypothesis, $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_b, \boldsymbol{\Sigma}_0)$ resulting in an expected Chi-Squared distribution for the first term in (4.21) [29],

$$(\mathbf{x} - \boldsymbol{\mu}_b)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) | H_0 \sim \chi_K^2, \tag{4.22}$$

which has an expected value of K. According to (4.17) the second term's expected value follows,

$$-E\big[\mathbf{x}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{x} \big| H_0\big] = -(Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0) + \boldsymbol{\mu}_b^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b). \tag{4.23}$$

With the remaining terms having expected values given by,

$$E\big[2(\mathbf{s}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{x}) - \mathbf{s}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{s} | H_0\big] = 2(\mathbf{s}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b) - \mathbf{s}^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} \mathbf{s}, \tag{4.24}$$

which results in an overall expected value given by,

$$\mu_{H_0} = E\big[l(\mathbf{x}) \big| H_0\big] = K - Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0) - (\boldsymbol{\mu}_b - \mathbf{s})^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{\mu}_b - \mathbf{s}). \tag{4.25}$$

Next, the variance is calculated according to,

$$\sigma_{H_0}^2 = E\Big[\big((\mathbf{x} - \boldsymbol{\mu}_b)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) - (\mathbf{x} - \mathbf{s})^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}) - \mu_{H_0}\big)^2 \Big| H_0\Big]. \tag{4.26}$$

When μ_{H_0} is substituted back in the result is given by,

$$\sigma_{H_0}^2 = E\Big[\big((\mathbf{x} - \boldsymbol{\mu}_b)^{\mathrm{T}} \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) - (\mathbf{x} - \mathbf{s})^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}) + (\boldsymbol{\mu}_b - \mathbf{s})^{\mathrm{T}} \boldsymbol{\Sigma}_1^{-1} (\boldsymbol{\mu}_b - \mathbf{s}) + Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0) - K\big)^2 \Big| H_0\Big]. \tag{4.27}$$

The square is distributed and cancellations performed. The result is a summation of expectations given by,

$$\begin{aligned}\sigma^2_{H_0} = E\Big[\big(\mathbf{x}^{\mathrm{T}}(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_0\Big] - 4E\Big[\big(\mathbf{x}^{\mathrm{T}}(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big] \ldots \\ + 4E\Big[\big(\mathbf{x}^{\mathrm{T}}(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\mathbf{s}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_0\Big] + 2W_0E\Big[\mathbf{x}^{\mathrm{T}}(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big] \ldots \\ + 4E\Big[(\boldsymbol{\mu}_b^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}\mathbf{x})^2\Big|H_0\Big] - 8E\Big[(\boldsymbol{\mu}_b^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\mathbf{s}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_0\Big] - 4W_0\Big[\boldsymbol{\mu}_b^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_0\Big] \ldots \\ + 4E\Big[(\mathbf{s}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\mathbf{x})^2\Big|H_0\Big] + 4W_0E\Big[\mathbf{s}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_0\Big] + W_0^2, \quad (4.28)\end{aligned}$$

where W_0 is a constant with respect to $\mathbf{x}$ defined by,

$$W_0 = \boldsymbol{\mu}_b^{\mathrm{T}}(\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b - 2(\boldsymbol{\mu}_b^{\mathrm{T}}\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - K. \quad (4.29)$$

Each expectation is calculated individually and after cancellation of terms and simplification the variance of the detection statistic under H_0 is found to be,

$$\sigma^2_{H_0} = 2Tr\Big(\big((\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_0^{-1})\boldsymbol{\Sigma}_0\big)^2\Big) + 4\big((\mathbf{s} - \boldsymbol{\mu}_b)\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}(\mathbf{s} - \boldsymbol{\mu}_b)\big). \quad (4.30)$$

The mean and variance under H_1 are calculated in a similar fashion resulting in,

$$\mu_{H_1} = (\mathbf{s} - \boldsymbol{\mu}_b)^{\mathrm{T}}\boldsymbol{\Sigma}_0^{-1}(\mathbf{s} - \boldsymbol{\mu}_b) + Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) - K, \quad (4.31)$$

and

$$\sigma^2_{H_1} = 2Tr\Big(\big((\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_1\big)^2\Big) + 4\big((\mathbf{s} - \boldsymbol{\mu}_b)\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}(\mathbf{s} - \boldsymbol{\mu}_b)\big). \quad (4.32)$$

More detailed derivations of μ_{H_0}, $\sigma^2_{H_0}$, μ_{H_1}, and $\sigma^2_{H_1}$ are detailed in Appendix A. The accuracy of these parameters have been verified through numeric simulation.

Each of these results are dependent upon $\boldsymbol{\Sigma}_0$. When applied for the signal-independent detection statistic (4.16), $\boldsymbol{\Sigma}_0$ in (4.25)-(4.32) is replaced with $\hat{\boldsymbol{\Sigma}}_0$. In this case, the mean and variance do not depend on $\mathbf{x}$ and the resultant conditional PDF is stationary. On the other hand, in the signal-dependent model $\boldsymbol{\Sigma}_0$ varies as a function of the observed pixel. In turn, this means that *the mean*

and variance of the detection statistic are both dependent upon the observed pixel. This result suggests that if signal-dependent noise is significant, the true FAR may vary from the designed FAR under the stationary assumption.

In the following chapter, performance is measured using three separate thresholding methods: signal-independent detection statistic (4.16) with a static threshold, signal-dependent detection statistic (4.15) with a static threshold, and signal-dependent detection statistic with a dynamic threshold. The signal-independent, static decision thresholds are calculated from the theoretical Normal distributions defined by (4.25) and (4.30) when $\mathbf{\Sigma}_0 = \hat{\mathbf{\Sigma}}_0$. Signal-dependent, static thresholds are calculated in the same manner, resulting in identical decision thresholds between these two methods. However, one would expect to see these identical thresholds result in different FARs due to differences in the distributions between the signal-independent and dependent results. It is reasonable to assume that in the signal-dependent case, a static threshold is sub-optimal since its behavior is not modeled by a single distribution. For the signal-dependent case, dynamic threshold case each threshold is calculated according to the Normal distribution defined by (4.25) and (4.30) where $\mathbf{\Sigma}_0$ is approximated by $\hat{\mathbf{\Sigma}}_b + diag(\boldsymbol{\gamma} \odot \tilde{\mathbf{b}} + \boldsymbol{\beta})$ for simulated data and $\hat{\mathbf{\Sigma}}_b + diag(\boldsymbol{\gamma} \odot \mathbf{x} + \boldsymbol{\beta})$ for real data.

CHAPTER V

RESULTS

The chapter steps through all of the results of the system model testing as well as the detection statistic performance. The application of the system transfer function is described and a new variable is introduced before the analysis of results. Additionally, experiments are designed and carried out on the datacubes of interest in order to investigate the effects of signal-dependent noise on hyperspectral target detection performance. The testing methodology for both the simulated and experimental results are outlined before the performance analysis.

5.1 System Modeling

This section showcases the performance of the system transfer function by estimating the raw data output of the sensor from the calibrated radiance measurements. The application of the transfer function is explained before being applied to integrating sphere data as a baseline performance test. After some tuning, the system model is applied to each target and checked for accuracy.

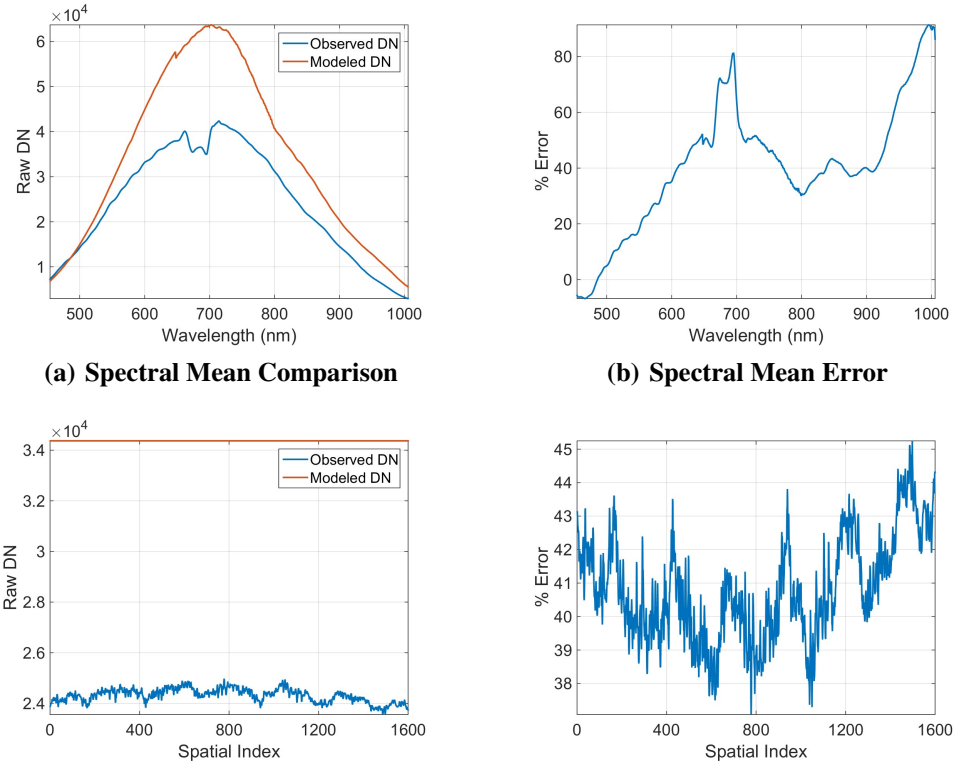
5.1.1 Testing Methodology

After outlining the design of the HSI system in Chapter III, the system transfer function term (3.1) is used to predict the raw data readout of the FPA for a set of known radiance values. Testing begins with some proof of concept work using data from an integrating sphere. The use of this highly

characterized illumination source allows for in depth analysis of the accuracy of the system model before being applied to spatially varying scene data. These tests utilize the following equation,

$$\hat{N}(i,k) = T(i,k) \times L(i,k) + N_1(i,k), \tag{5.1}$$

where $\hat{N}(i,k)$ is the modeled raw data value, $T(i,k)$ is the system model's transfer function calculated in (3.1), and $L(i,k)$ is the observed spectral radiance which is known when using the integrating sphere data. Figure 5.1 compares spectral and spatial averages for this modeled and observed data. For these results, the foreoptic was removed to virtually eliminate vignetting in the spatial



(a) Spectral Mean Comparison **(b) Spectral Mean Error**

(c) Spatial Mean Comparison **(d) Spatial Mean Error**

Figure 5.1: Untuned System Modeling Results. These results come from integrating sphere data. According to the error plots the modeled data is 41% higher than the observed. The spectral shape of error is consistent with all other results.

dimension. These results show that the modeled data is much higher than the observed – by roughly

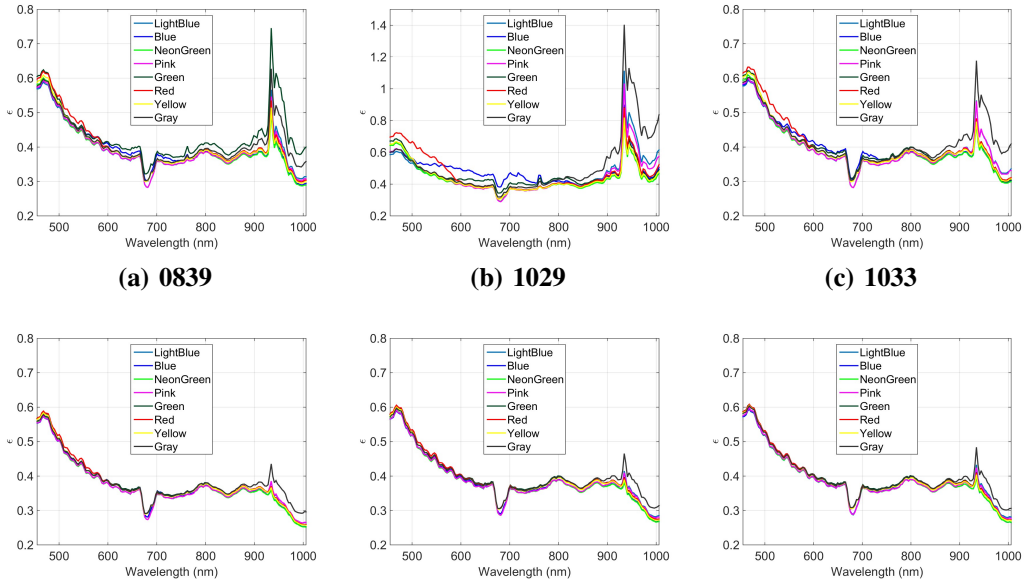
41%. The error plots show a relatively constant error in the spatial dimension while the spectral error deviates from this 40% in the 400-600 and 900-1000nm ranges. The results of other datacubes collected in this same experiment display the same spectral shape, but the overall average error changes drastically from 41% to 84%. These results are included in Appendix B.

The source of this spectral error is believed to come from inaccurate transmission / reflectance / efficiency specifications for one or more of the system's components. The contribution from the foreoptic system or UV filter is believed to be minimal because the spectral error shape does not change when these elements are removed. However, the fact that the overall mean error changes from cube to cube suggests that another factor is at play. This could be a result of mismatches between the true f/# of the lens and the value used in the model. The f/# of the lens is controlled by hand, introducing a possible source of inaccuracy. At this point, these are theories, and there is the possibility that the system model being used is overlooking an element with significant contribution to the system's performance.

To account for this error a tuning parameter $\epsilon(i,k)$ is added to 3.1 resulting in the following,

$$N(i,k) = L(i,k)\left(g(i,k)\frac{\pi}{4(f/\#)^2+1}K(i,k)\eta_{QE}(\lambda)\tau(\lambda)\frac{\lambda}{hc}t_d\Delta\lambda A_d\right)\epsilon(i,k)+N_1(i,k). \quad (5.2)$$

The precise source of $\epsilon(i,k)$ is not known making it difficult to model. For this report, $\epsilon(i,k)$ is assumed to be spatially invariant, i.e. $\epsilon(i,k) \approx \epsilon(k)$, and its spectral contents are calculated using the scene data. Estimates of the parameter are calculated for each datacube using the calibrated panel data and are represented in Figure 5.2. The 0839, 1029, and 1033 results show that these estimates change measurably from panel to panel. In general, the results are consistent for spectral bands less than 900nm, but performance suffers at longer wavelengths. These inconsistencies are believed to be due to the low signal level – brought on by the solar positioning during the 0839 collection and the short exposure times of the 1029 and 1033 datacubes – combined with the low



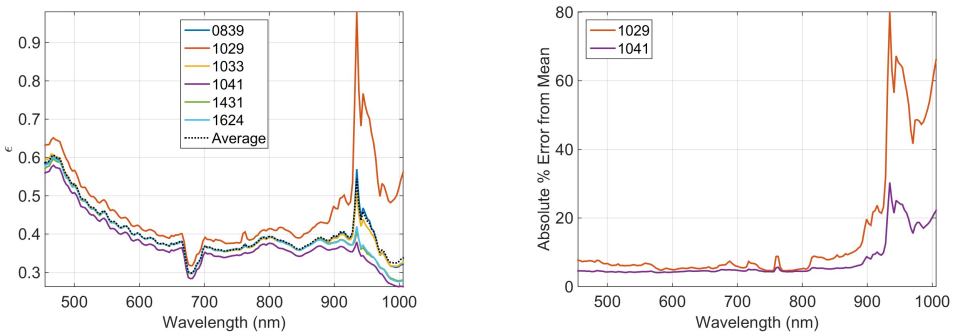
(a) 0839 **(b) 1029** **(c) 1033**

(d) 1041 **(e) 1431** **(f) 1624**

Figure 5.2: Tuning Parameter : All Results. Note that the 1029 results are presented at a different scale than all of the other results.

detector response in this spectral range. The 1041, 1431, and 1624 estimates do not display a high degree of variability from panel to panel due to the higher overall signal levels.

The mean tuning parameter is calculated for each datacube and the results are shown in Figure 5.3(a). Again, results suggest that the tuning parameter is fairly consistent before 900nm. Figure 5.3(b) shows the percent error for the highest and lowest tuning parameters with respect to the mean. The average error for wavelengths less than 900nm is 6.81% for the 1029 data and 4.76% for the 1041 cube. For wavelengths above 900nm this jumps to 48.32% and 18.06%, respectively. The increase in error is theorized to be due to the low signal levels leading to increased degrees of uncertainty in this spectral region. This limited data set seems to indicate that the exposure time plays a role in the value of this tuning parameter. The average tuning parameter $\bar{\epsilon}$ – represented by the dotted line in Figure 5.3(a) – is used for all of the datacubes in the assessment of the accuracy of the system model.



(a) Tuning Parameters **(b) Error**

Figure 5.3: Tuning Parameter : Datacube Mean & Error. Plots show the calculated tuning parameters ϵ for each datacube along with the overall average. The 1029 and 1041 results display the largest deviations from the mean and this error is represented as well.

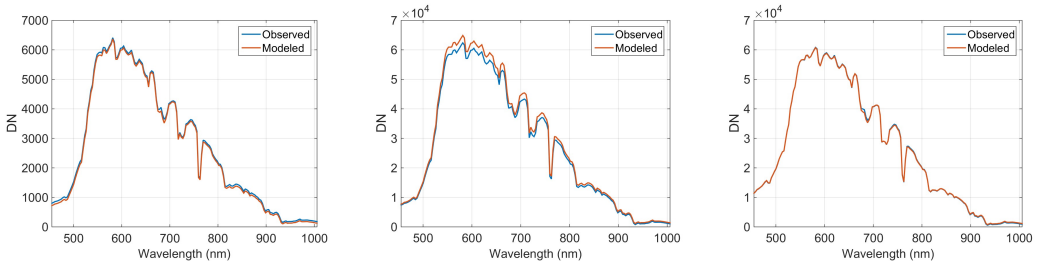
5.1.2 Results

Using this newly defined $\bar{\epsilon}$, the refined system model (5.2) is used to predict the raw data results of the six datacubes. The mean spectrum of each target panel is used as the reference radiance level $L(\lambda)$. Equation (5.2) is applied to estimate the mean raw data spectrum. This estimate is then compared to the observed raw data. Sample results from this work are contained in Figure 5.4. This figure portrays the accuracy of the modeling technique for the yellow and green panels in the 1029, 1041, and 1624 datacubes.

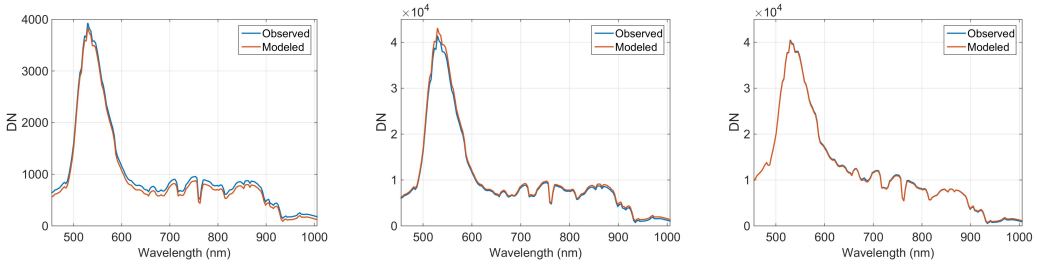
The mean absolute percent error is calculated for each result according to the following formula,

$$\mathcal{E} = \frac{1}{K} \sum_{k=1}^{K} \frac{\left| N(k) - \hat{N}(k) \right|}{N(k)}. \tag{5.3}$$

Of all six results in Figure 5.4 the green spectra from the 1029 collection shows the highest error at 12% while the 1624 collection has the two lowest error metrics, both slightly below 4%. This is not surprising when recalling Figure 5.3 as the ϵ of the 1624 collection is closer to $\bar{\epsilon}$ than the other two datacubes. Figure 5.5 compares error $\mathcal{E}$ for all of the data under test. The results show that the system transfer function in (5.2) is able to accurately model the raw camera data. In general, the



(a) 1029 : Yellow (b) 1041 : Yellow (c) 1624 : Yellow



(d) 1029 : Green (e) 1041 : Green (f) 1624 : Green

Figure 5.4: System Model Results : Spectra Comparison. Each of these results show that the modeled spectra is accurately predicting the observed raw data. Even the two datacubes with the highest degree of error (1029 and 1041) seem to be modeling accurately. Appendix B contains these spectral comparisons for all targets, broken down by datacube.

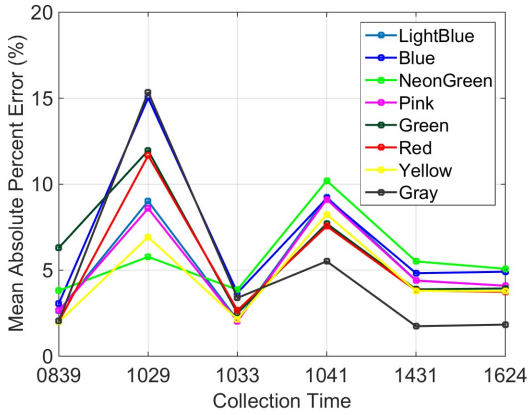


Figure 5.5: System Model Results : Error Trends. These results suggest that the model performance is more of a function of data cube than the target under test. With the exception of the 1029 datacube, all data is at 10% error or below. See Appendix B for the table containing the values behind this figure.

results are within 10% and are therefore considered to be accurate models of the observed spectra. The 1029 data records the highest errors with the gray and blue panels showing 15% error while the 1041 results are the next worst performers. The 0839, 1431, and 1624 results all display similar performance. These results support the theory that the tuning parameter is – to some degree – a function of exposure time. Further tests will need to be conducted in order to thoroughly investigate this and to attempt to provide a method of predicting ϵ. This could be done by collecting several datacubes with varying exposure times while keeping all other factors constant. These results could be used to find the relationship between ϵ and exposure time.

This sensor system transmission model is just one element of this research. These results display this model's ability to predict real work performance, but all of these results are representative of the mean target spectra. Section 3.2 explains the physical processes that account for sensor noise variations. The ability to accurately model this sensor noise is vital to creating an end-to-end HSI system performance model. The following sections investigate the effects of this sensor noise on target detection performance.

5.2 Detection Statistic Performance with Simulated Background

This section investigates the effects of signal-dependent noise on target detection performance through the implementation of a simulated background. The experimental design is outlined before the presentation of results. In the results portion, the Normality of the detection statistic distributions is evaluated in addition to the subsequent false alarm rates.

5.2.1 Data Review and Testing Methodology

With the definition of the signal-dependent detection statistic in place (4.15), a method of performance analysis is created with the goal of assessing both the effects of incorporating this signal-dependent noise and the ability to predict the performance of the detection statistics regardless of the noise model implemented. Analysis begins with the design of simulations that investigate the distributions of (4.15) and (4.16) for each target and datacube. Simulated backgrounds are generated according to,

$$\mathbf{b} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_b, \hat{\boldsymbol{\Sigma}}_b), \tag{5.4}$$

and use the statistics from the real data in an effort to produce realistic simulation results. $\hat{\boldsymbol{\Sigma}}_b$ is calculated for each datacube – according to (4.13) – as is the sample background mean $\hat{\boldsymbol{\mu}}_b$. Simulations are designed to generate 200,000 noiseless pixels in every iteration. Noise is added to each realization according to the heteroskedastic model (3.14). This results in the distribution given by,

$$\mathbf{x} \sim \mathcal{N}\big(\hat{\boldsymbol{\mu}}_b, \hat{\boldsymbol{\Sigma}}_b + diag(\boldsymbol{\gamma} \odot \tilde{\mathbf{b}} + \boldsymbol{\beta})\big). \tag{5.5}$$

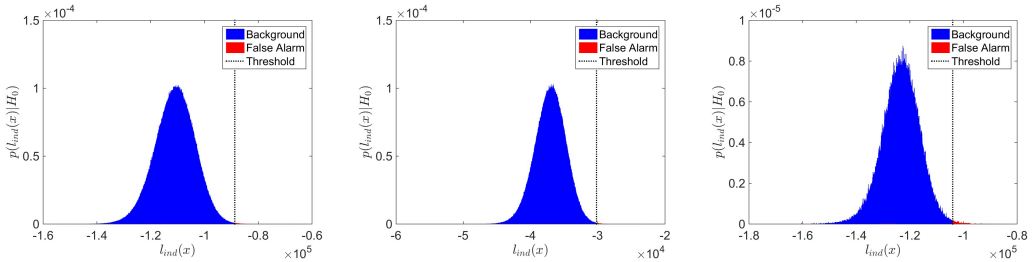
Note that each noisy observation $\mathbf{x}$ is a *single realization* of a Normally distributed random process whose variance is a function of $\tilde{\mathbf{b}}$. Each $\mathbf{x}$ is then individually input into the signal-dependent (4.15) and signal-independent detection statistics (4.16). When this has been completed for each $\mathbf{x}$, a histogram of the result is calculated. This histogram acts as an approximation to $p(l(\mathbf{x})|H_0)$. In Section (4.4), it was stated that these results are assumed to be Normally distributed due to the

central limit theorem. Thresholds are calculated according to the Normal distribution governed by (4.25) and (4.30). For the signal-independent detection statistic results (4.16), a single τ is calculated for the distribution. The signal-dependent (4.15) results are thresholded using two methods: static and dynamic. The static threshold is a single τ used to threshold the distribution results. The dynamic threshold calculates a new τ for each $\mathbf{x}$ resulting in a varying threshold. This method is implemented in an effort to address the observation in Section 4.4 that the expected distribution of the signal-dependent detection statistic changes with each new $\mathbf{x}$. For all simulations, τ is designed to produce a 10^{-3} FAR with the theoretical Normal distributions. The three thresholding methods are applied and FARs are calculated. This process is repeated 100 times in Monte Carlo fashion and averaged in an effort to arrive at a more accurate estimate of the FAR behavior. This simulation strategy is applied to all eight targets for all six datacubes.

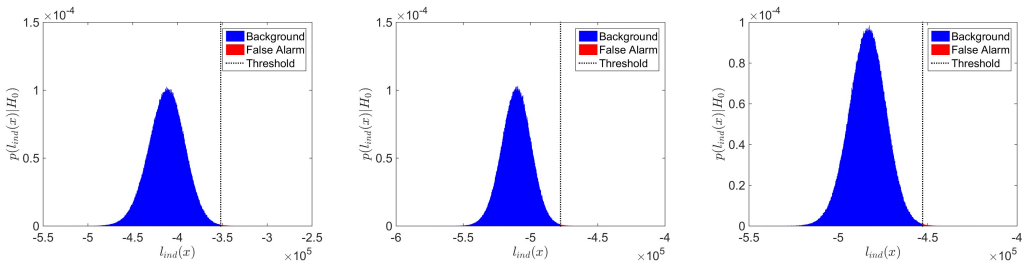
The following section looks at how well the theoretical Normal distributions governed by (4.25) and (4.30) model the distributions of the (4.15) and (4.16) with a simulated background input.

5.2.2 Distribution Normality

This section highlights the changes that occur to the Normality of the distributions as a function of noise model, target, and data cube. It begins with the analysis of two targets, the green and yellow panels. Masks are used to calculate the mean spectra of each panel $\mathbf{s}$. The signal-independent and signal-dependent detection statistic distributions for these two panels can be found in Figures 5.6-5.9. Figures 5.6 and 5.8 show the signal-independent results while 5.7 and 5.9 are representative of the signal-dependent results. In each of these figures, the decision threshold τ is represented by the black dotted line. Pixels correctly classified as background are blue while false alarms are red.

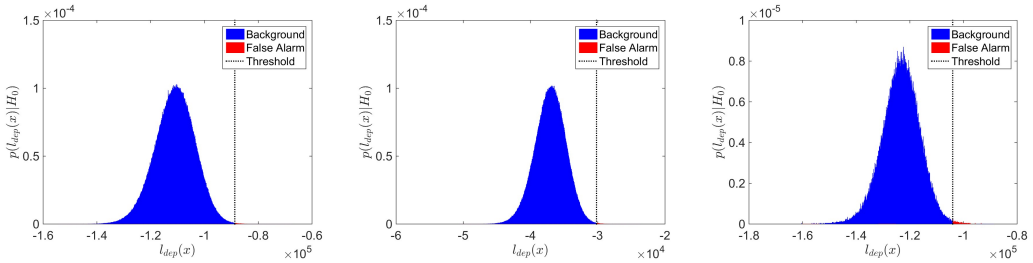


(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

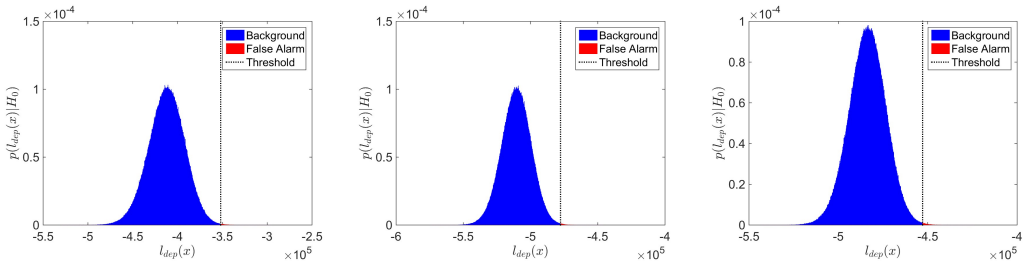


(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure 5.6: Simulated Distribution Results : Yellow, Signal-Independent. Each distribution looks to be fairly symmetrical and approximately Normal. From a strictly visual perspective these results show no major variation in Normality from one datacube to the next. Results are reflective of a goal CFAR of 10^{-3}.

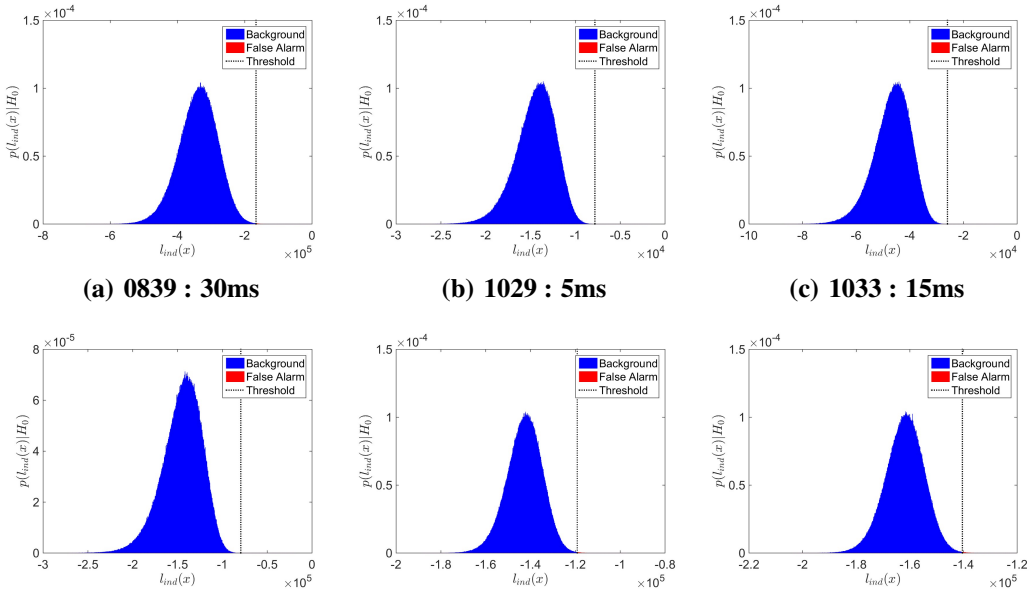


(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**



(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

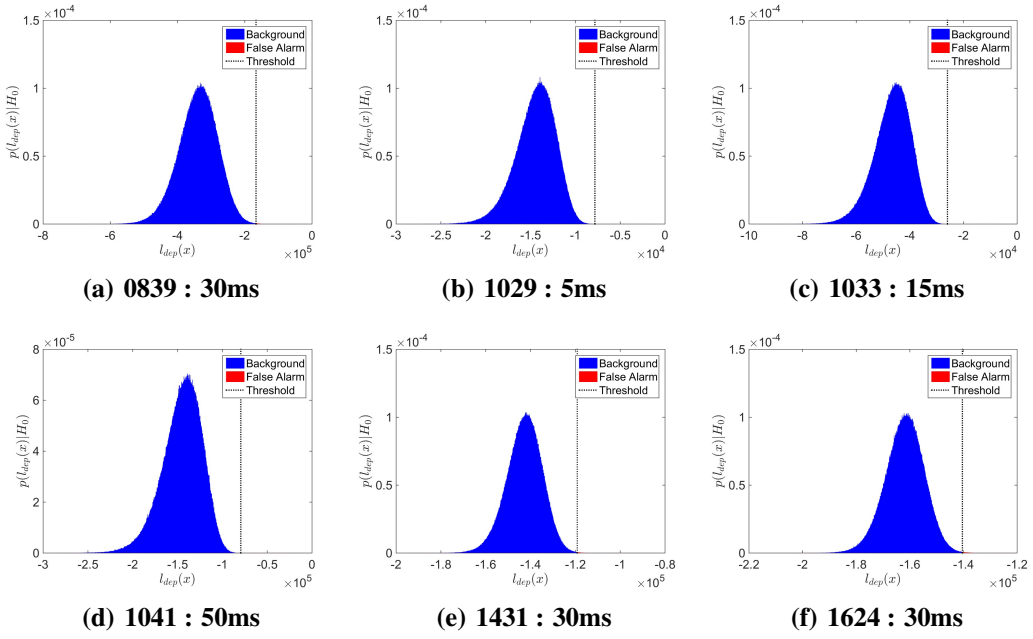
Figure 5.7: Simulated Distribution Results : Yellow, Signal-Dependent. Distributions look identical to Figure 5.6. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure 5.8: Simulated Distribution Results : Green, Signal-Independent. In contrast to the yellow panel results in Figures 5.6 and 5.7, these distributions display varying degrees of Normality. The distributions resulting from the 0839-1041 collections all display a left handed tail. While the remaining two look more symmetric – and therefore closer to Normal – than the others, but careful inspection reveals that these two display a slight left handed tail. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure 5.9: Simulated Distribution Results : Green, Signal-Dependent. Distributions look identical to Figure 5.8. Results are reflective of a goal CFAR of 10^{-3}.

The observation can be made from Figures 5.6-5.9 that distribution changes as a function of the noise model are difficult to see. The distribution characteristics change as a function of datacube and target under test, but there are no observable differences as a function of the detection statistic itself. Table 5.1 shows that the distributions are effectively independent of the noise model used. To illustrate this point, the analysis is expanded to the rest of the targets in Figure 5.10 investigating the change in the Normality of the distributions that occur as a function of the noise model being used. Here, the normalized difference is defined as the difference between the signal-independent and signal-dependent measures divided by the signal-dependent. Results suggest that the conclusion drawn from Table 5.1 may be applied to the entire data set. *There is little to no effective change in the distribution due to the incorporation of signal-dependent noise for these simulated backgrounds.* While gross distribution changes are not observed, there may be differences in the extreme tails of

		Mean	Variance	Skew	Kurtosis
0839	**Yellow - Independent**	-1.10979e+05	5.18879e+07	-0.15090	0.04702
	Yellow - Dependent	-1.10979e+05	5.18878e+07	-0.15090	0.04702
	Green - Independent	-3.38746e+05	3.10364e+09	-0.25792	0.09001
	Green - Dependent	-3.38746e+05	3.10364e+09	-0.25792	0.09001
1029	**Yellow - Independent**	-3.70635e+04	4.95381e+06	-0.13606	0.02963
	Yellow - Dependent	-3.70635e+04	4.95382e+06	-0.13606	0.02963
	Green - Independent	-1.44269e+04	4.57211e+06	-0.50766	0.37990
	Green - Dependent	-1.44269e+04	4.57212e+06	-0.50767	0.37990
1624	**Yellow - Independent**	-4.83761e+05	9.98075e+07	-0.07951	0.01358
	Yellow - Dependent	-4.83761e+05	9.98073e+07	-0.07951	0.01358
	Green - Independent	-1.61897e+05	4.83608e+07	-0.18651	0.10016
	Green - Dependent	-1.61897e+05	4.83608e+07	-0.18651	0.10015

Table 5.1: Simulated Distribution Comparison : Yellow and Green Panels. These results confirm the qualitative conclusions drawn from Figures 5.6-5.9. The signal-dependent and signal-independent detection statistic distributions are effectively equivalent. This table represents two of eight targets in three of six data cubes, but these results are representative of the data as a whole.

the distributions. The changes in the distribution tails can impact FAR performance, especially in low FAR applications such as this.

In high SNR data, the fact that noise makes up a small percentage of the observed data implies that the accuracy of the noise model may not greatly contribute to the accuracy of the algorithm. On the other hand, in low SNR data the noise makes up a larger portion of the observation. In this case accurate noise modeling is expected to have beneficial effects to performance while inaccurate modeling could be detrimental. Therefore, it is expected that lower SNR data cubes will be more influenced by the noise model than the higher SNR sets. This hypothesis explains the trend in the skew of the red panel between the 1029 and 1041 collections. The 1029 cube, having an exposure time of 5ms is the most noisy cube of the six. The 1033 and 1041 cubes each have a longer exposure time, effectively increasing the SNR. While the result fits with expected performance, the trend is only seen in one of the eight targets.

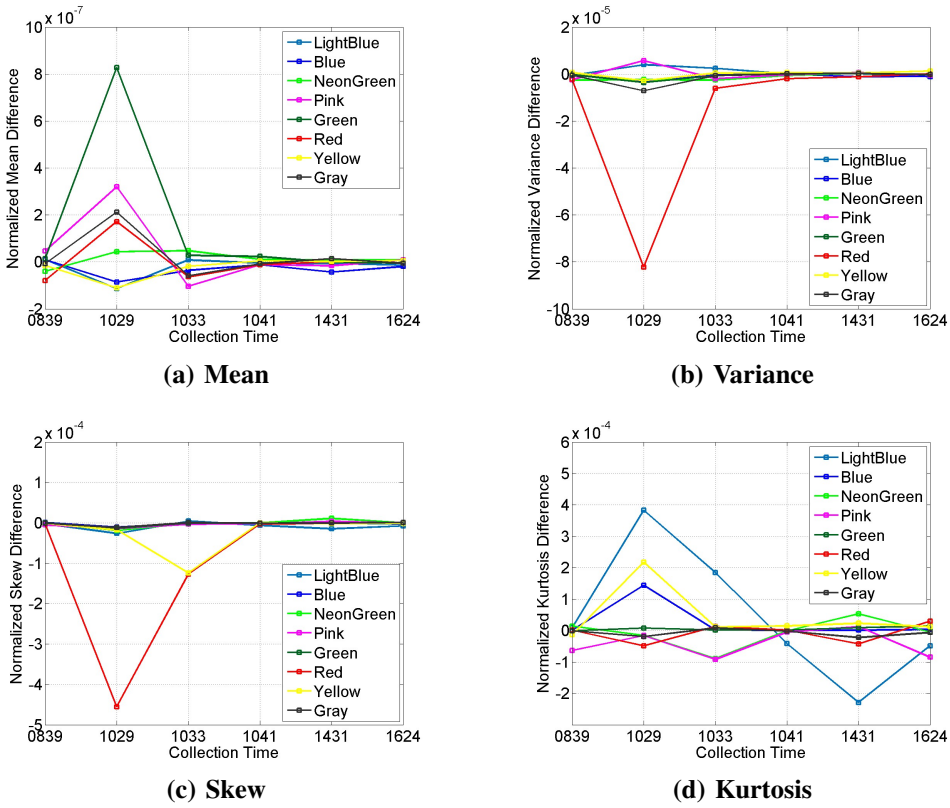
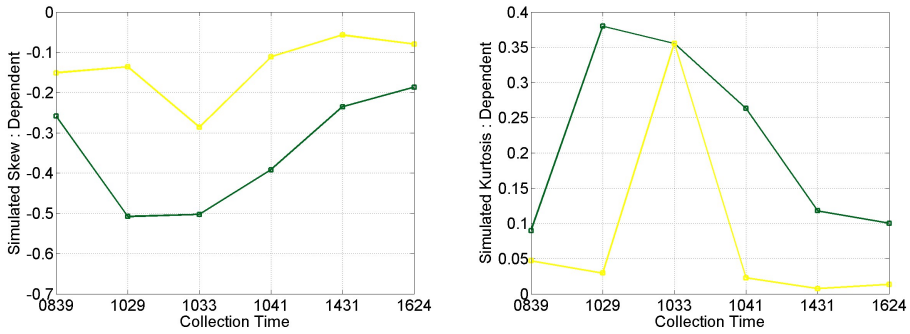


Figure 5.10: Effect of Signal-Dependent Noise on Distributions of Simulated Detection Statistics. The magnitude of all values is small enough that the signal-independent and signal-dependent distributions are considered to be equal to one another.

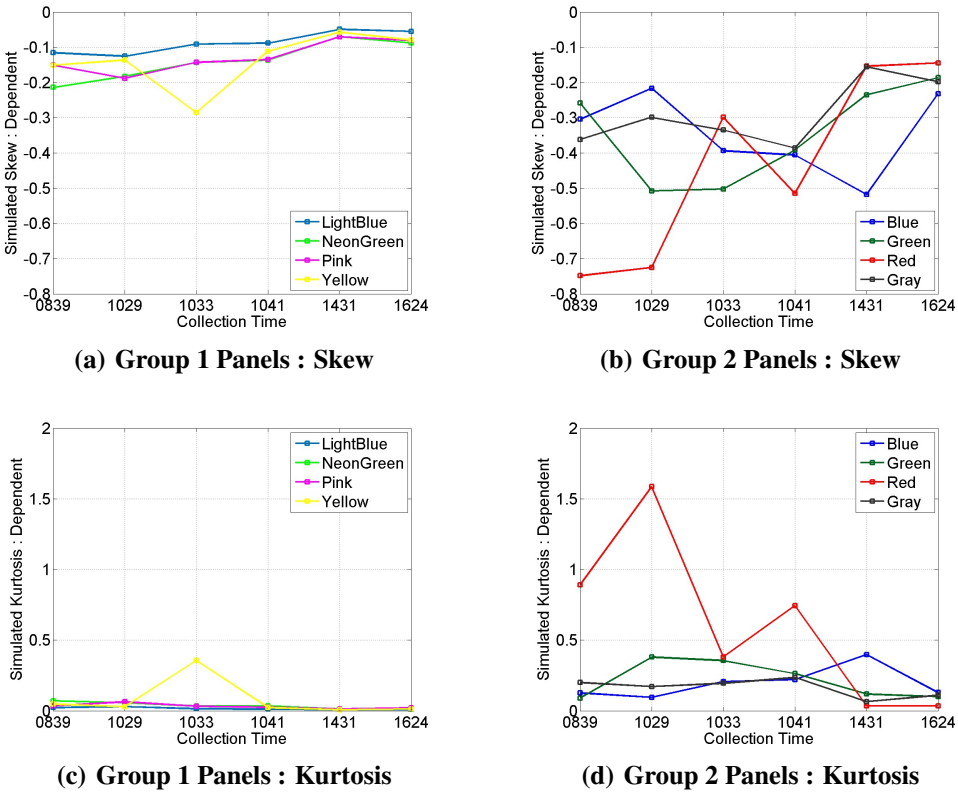
While the detection statistic distributions do not change much as a function of the underlying noise model, they vary greatly as a function of target panel (s) and collection time. The yellow panel results in Figures 5.6 and 5.7 seem to behave according to the Normal assumption made in Section 4.4, while the green distributions in 5.8 and 5.9 clearly display varying degrees of asymmetry. Figure 5.11 investigates the Normality of these two panels in depth and quantifies the clear differences in Normality observed in Figures 5.6-5.9. These results indicate that both panels produce skewed statistics, but also confirm the observation that the green panel exhibits a higher degree of skew. A comparison of the kurtosis measure by Figure 5.11(b) also indicates that the yellow panel results are closer to approximating a Normal distribution than the green.



(a) Skew **(b) Kurtosis**

Figure 5.11: Simulated Normality Results : Yellow and Green Panels. Plot attempts to visualize the Normality trends of Figures 5.6-5.9. A skew measure of zero indicates a symmetric distribution. These results confirm the visual observation that the green panel distributions display a heavier left handed tail than the yellow panel distributions. This is indicated by the negative skew measure.

These Normality observations are extended to the other panels in the scene and the targets are separated into two groups for analysis. Once again, the results are representative of the detection statistic for the background pixels. The target signal $\mathbf{s}$ changes for each panel. Group 1 consists of the four panels whose detection statistic distributions fit the Normality assumption most closely while Group 2 consists of the remaining four panels. Figure 5.12 shows the Normality measures of all panels after separating the results into the two groups. Group 1 consists of the light blue, neon green, pink and yellow panels. Results from this group show near-zero values for skew and kurtosis and appear to be fairly robust to solar illumination levels and exposure time. The Group 2 panels – blue, green, red, and gray – each show results of a higher magnitude, suggesting non-Normal characteristics, but also show a high degree of variability from collection to collection. In some cases the Normal approximation may be applicable, but in most cases these four panels behave in a non-Normal manner. *These results indicate that the validity of our normality assumptions changes with the target under test.*



(a) Group 1 Panels : Skew

(b) Group 2 Panels : Skew

(c) Group 1 Panels : Kurtosis

(d) Group 2 Panels : Kurtosis

Figure 5.12: Simulated Normality Results : All Panels. Results suggest that half of the panels–light blue, neon green, pink, and yellow–have detection statistic distributions that behave in a near-Normal fashion. The remaining four have more asymmetric distributions. These results are representative of the signal-dependent detection statistic distribution and are nearly identical for the signal-independent distributions. See Appendix C for tables detailing the measures of Normality for each signal model.

The results in Figures 5.11-5.12 are representative of the signal-dependent results. Signal-independent results are not addressed directly because, according to Figure 5.10, their distribution characteristics are considered identical to the signal-dependent results. Reference Appendix C for tables contained the values of the data behind these plots.

5.2.3 False Alarm Performance

While the previous section shows that the incorporation of signal-dependent noise into the target detection model has no major effect on the overall shape of the resultant distribution, it is necessary to investigate whether or not the noise model affects the FAR performance in any manner. The previous section's analysis may overlook slight changes in the extreme distribution tails. The behavior in these tails drives FAR performance, especially for low FAR applications. Of particular interest is the performance change between the static and dynamic thresholding methods for the signal-dependent results. Analysis begins with the green and yellow panels in Figure 5.13 before expanding to the entire dataset. These results show the green panel having a lower FAR than the

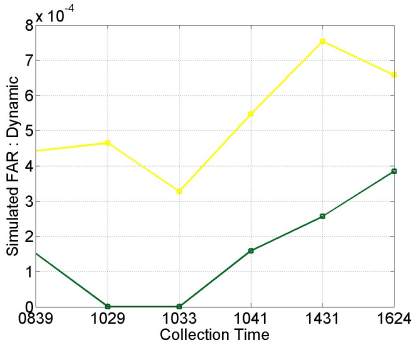
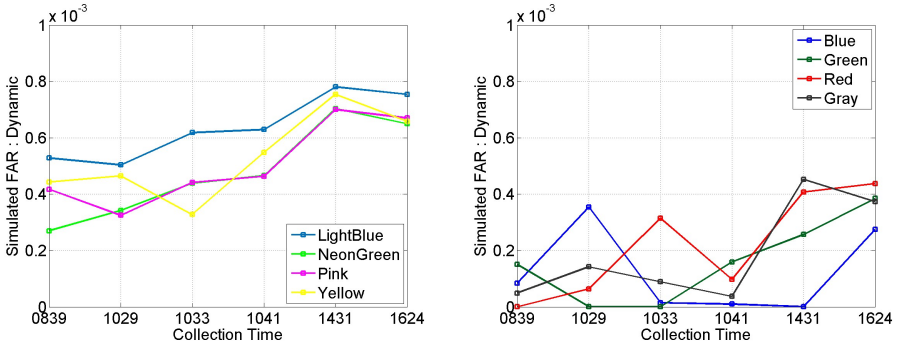


Figure 5.13: Simulated FAR Results : Green and Yellow Panels. Plot attempts to visualize the FAR trends of Figures 5.7 and 5.9 using the dynamic thresholding method. The green panel returns a lower FAR for all datacubes. The τ used for thresholding is calculated for a CFAR of 10^{-3}

yellow results in all of the tested datacubes despite τ set for constant FAR. The previous analysis showing that the green panel is heavily skewed may lead one to believe that its FAR would suffer as a result. Instead, the heavy left-handed skew has dramatically decreased the FAR below the designed value of 10^{-3}. The fact that the yellow panel has a higher FAR for all examples is not inherently indicative of poorer performance. The yellow results are closer to Normal than the green

(Figure 5.11). Therefore, since the τ is calculated from a Normal distribution it is logical that the yellow results would perform closer to the designed 10^{-3} CFAR. Figure 5.14 expands this FAR analysis to all of the featured targets, separating them once again based off of their Normality. Due



(a) Group 1 Panels : FAR **(b) Group 2 Panels : FAR**

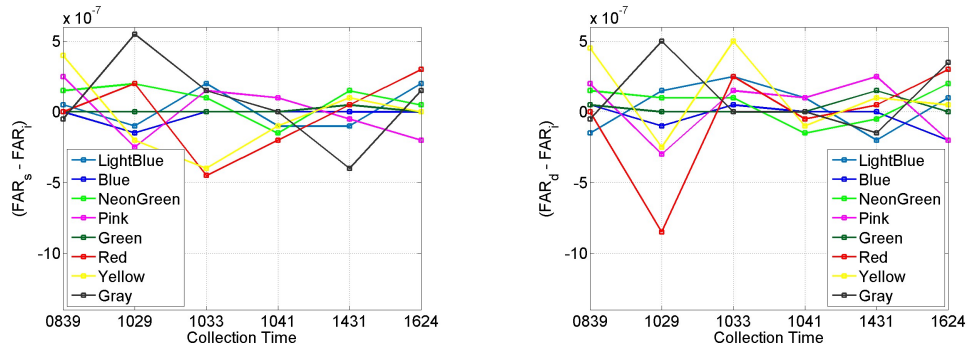
Figure 5.14: Simulated FAR Results : All Panels. Results suggest that the panels with with near-Normal behavior achieve FARs closer to the desired CFAR of 10^{-3} than the panels whose distributions have strong tails. Plotted results represent FARs when using the dynamic thresholding method. See Appendix C for tables containing data on all thresholding methods.

to the high degree of non-Normality in the results, comparing FAR is difficult. A simulated FAR near the designed CFAR indicates an accurately modeled distribution. However, while a FAR lower than the CFAR is an indicator of improved performance, it also reflects the inaccuracy of the Normality assumption. Instead of looking at FARs as a function of target and data set, this report will focus on the performance trends that occur as a function of the implemented noise model. Table 5.2 investigates the FAR of the yellow and green panels for all three thresholding methods. Similar to the measures of Normality in Section 5.2.2, these results suggest that the FAR may be independent of the thresholding method and noise model.

Figure 5.15 shows the changes in FAR performance due to the signal-dependent detection statistic using both thresholding methods while using the signal-independent FAR as a baseline metric. A

Table 5.2: Simulated FAR Comparison : Yellow and Green Panels. FAR of the yellow and green panel detection statistic distributions for the 0839, 1029, and 1624 datacubes. All three thresholding methods are represented. No consistent performance improvement is noted from any of the methods. Thresholds designed for a desired CFAR of 10^{-3}.

		Signal-Independent	Static	Dynamic
0839	**Yellow**	0.44225e-03	0.44265e-03	0.44270e-03
	Green	0.15070e-03	0.15070e-03	0.15075e-03
1029	**Yellow**	0.46545e-03	0.46525e-03	0.46520e-03
	Green	0.00075e-03	0.00075e-03	0.00075e-03
1624	**Yellow**	0.65825e-03	0.65825e-03	0.65830e-03
	Green	0.38500e-03	0.38500e-03	0.38500e-03



(a) Static FAR Difference **(b) Dynamic FAR Difference**

Figure 5.15: Effect of Signal-Dependent Noise on FAR of Simulated Detection Statistics. Plots reflect the performance changes due to the signal-dependent noise thresholding techniques using the signal-independent results as reference. Note that since a small FAR is desired, negative values indicate improved performance by the signal-dependent model. Results shown are for designed CFAR of 10^{-3}

data point at zero indicates that no performance change is introduced by the signal-dependent detection statistic, while a negative value shows improved performance. The largest observed improvement is the red panel in the 1029 collection, but it's an incredibly small improvement of roughly 8×10^{-7}. These results indicate once again that the noise model under test does not play a major role in the outcome. However, they do not compare the performance of the two signal-dependent thresholding methods. Figure 5.16 compares these methods. By showing the average FAR change for

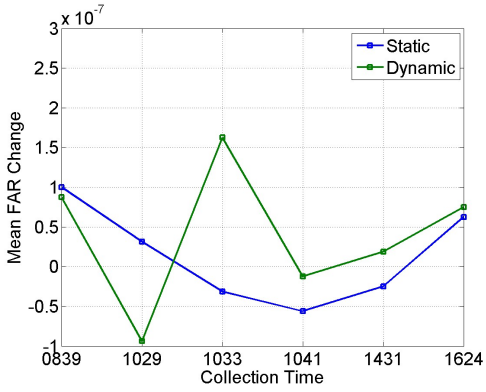


Figure 5.16: Simulated FAR Results : Average Change. The average FAR change attempts to represent how each data set performs overall. Results shown are for designed CFAR of 10^{-3}.

both signal-dependent thresholding methods and each datacube, the plots in Figure 5.2.3 attempt to communicate which of the two thresholding methods achieve overall better performance for a specific datacube. Similar to Figure 5.15, a mean FAR change less than zero indicates an average performance improvement over the signal-independent algorithm. Results show that the signal-independent detection statistic performs best for the 0839 and 1624 datacubes, the signal-dependent static threshold outperforms all others for the 1033 through 1431 collects, and the signal-dependent dynamic thresholding method performs best for the 1029 data set. It is important to point out that while these results indicate that one algorithm may outperform another, the individual performance changes for each target are minimal (Figure 5.15).

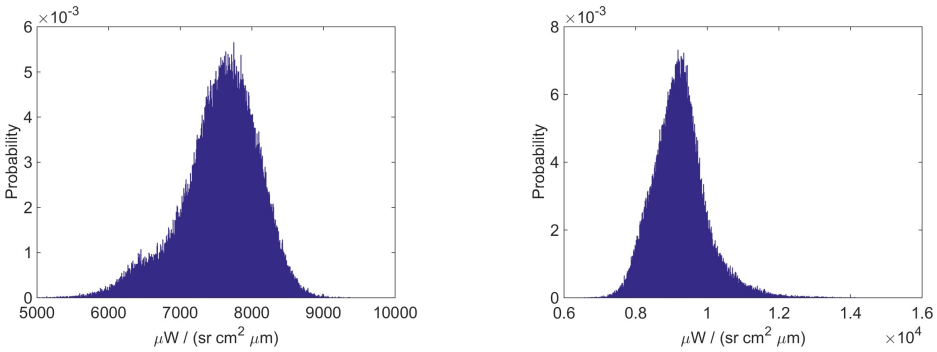
5.3 Experimental Results

The simulations investigate the performance of the detection statistics when presented with an ideal, Normally distributed background. However, real world data is typically non-Normal. This section investigates the Normality of the true background data, showcasing the differences between

it and the simulated backgrounds from Section 5.2. This data review is then followed by an analysis of the detection statistic performance with these backgrounds.

5.3.1 Data Review and Testing Methodology

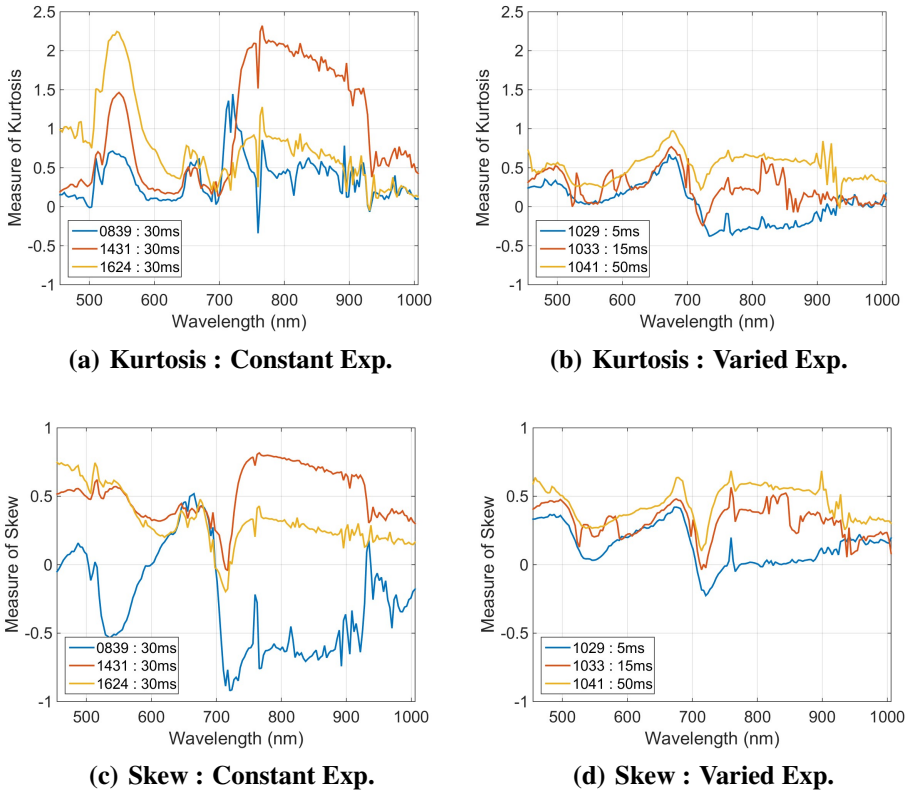
While the previous simulations assumed a Normally distributed background, Figures 5.17 and 5.18 show that the real background classes for these data sets do not follow this assumption. These



(a) 0839 **(b) 1431**

Figure 5.17: Example Background Distributions : 824nm. Histograms are representative of the background data in a single spectral channel. The results emphasize those present in Figure 5.18, that the distribution characteristics for a single band changes with each datacube.

plots look at the Normality of the background in each spectral band. Figure 5.17 hints at the non-Normality of the background data and suggests that distribution characteristics change with respect to datacube. This figure is only representative of the 824nm data. Figure 5.18 extends this by showing that the Normality is dependent upon wavelength and the datacube itself. For example, the background of the three cubes collected during the 1000 hour are much more similar than the other three cubes. Overall, skewness and kurtosis for these cubes are closer to zero meaning they are closer to a Normal distribution. The level of variation from cube to cube is smaller as well. This is due to a combination of constant solar angle, solar illumination, and nearly identical background



(a) Kurtosis : Constant Exp. **(b) Kurtosis : Varied Exp.**

(c) Skew : Constant Exp. **(d) Skew : Varied Exp.**

Figure 5.18: Normality of Background Class. This figure looks at the Normality of each band of the background class through the measures of skew and kurtosis. The backgrounds of the 1000 datacubes are both more Normal and more similar to each other than the three other cubes.

pixel sampling. Each of these cubes are able to use an identical background mask because the camera position was not changed between collections. This is evident in Figure 4.1.

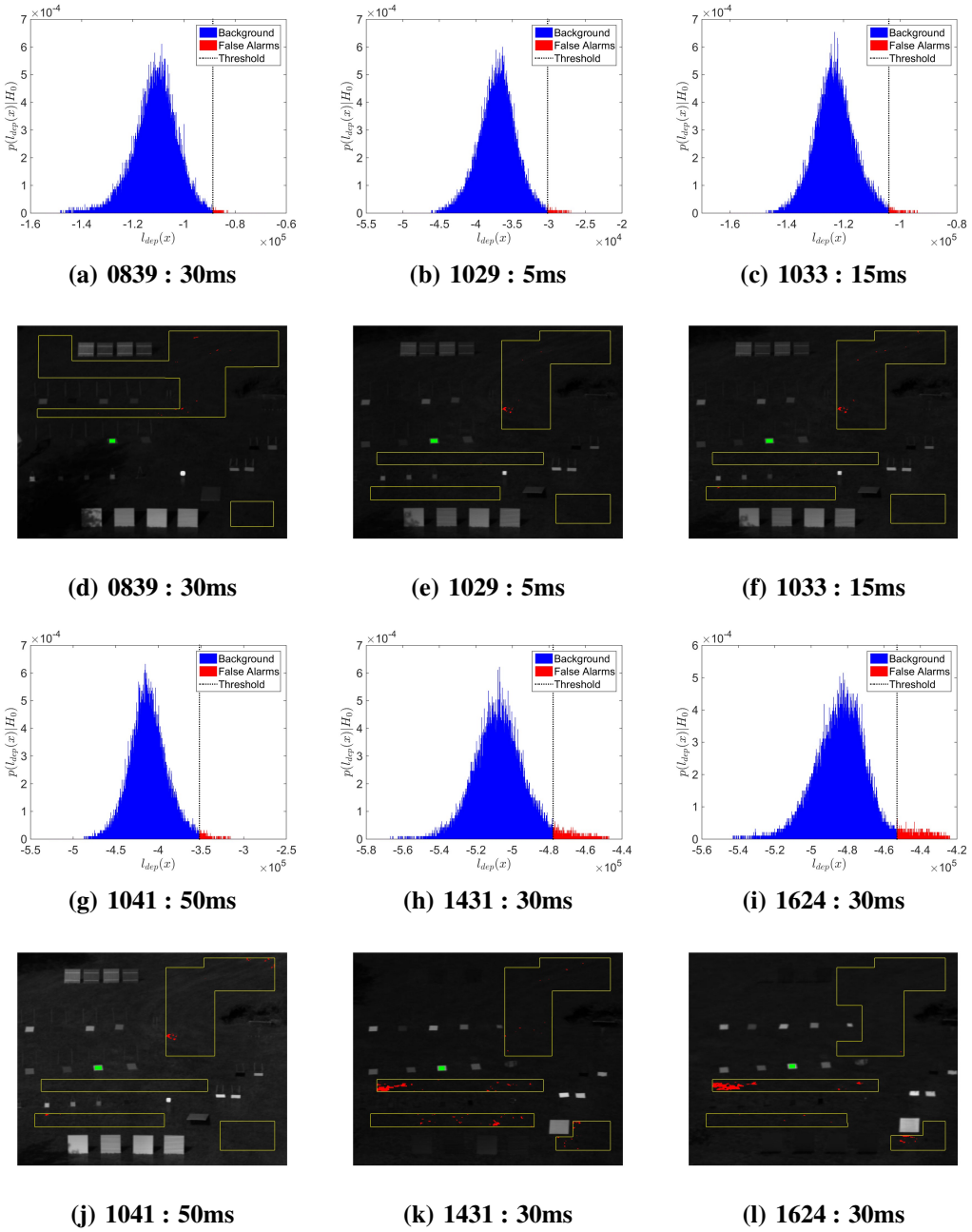
The pixel positioning of the man-made objects in the scene are slightly different for the constant exposure time targets due to changes in object locations as well as shifts in the camera base over the course of the two collection days. This results in the need for slight variations in the background mask from cube to cube. Additionally, the 0839 cube is experiencing direct shading across pixels that are typically used as the background class, resulting in a much different mask than the other

five cubes. Which likely contributes to the skew shape in Figure 5.18(c) that is vastly different than the other cubes.

The performance analysis continues with the experimental hyperspectral data by applying each background observation to the signal-independent and signal-dependent detection statistics. For this real data, the noiseless realization $\tilde{\mathbf{b}}$ is not known. Therefore the observed value $\mathbf{x}$ is used to approximate $\tilde{\mathbf{b}}$ in the signal-dependent LRT (4.15). This approximation is also used in the calculation of the mean (4.25) and variance (4.30) of the expected distributions through $\mathbf{\Sigma}_0$.

5.3.2 Detection Statistic Distribution Normality

Figures 5.19 and 5.20 show the experimental results of the application of the static threshold to the signal-dependent detection statistics for the yellow and green target panels. These results can be directly compared to their simulated counterparts in Figures 5.7 and 5.9. Unlike in Section 5.2, the signal-independent results are not shown as it has been made clear that, at least from a visual perspective, the distributions are indistinguishable. Grayscale images below each empirical pdf highlight the location of the false alarms. Recall from Section 5.2 that simulations with the yellow panel result in near-Normal distributions, and all six data cubes display a slight left skew. When those same panels are tested with real background data the results are visibly further from a true Normal distribution. The 1431 and 1624 sets show extensive right-handed tails leading to a high number of false alarms. Each of these distributions still look to be approximately Normally distributed. However, they are not as well behaved as the corresponding simulation results. The real background data seems to have an observable effect on the shape of the distributions. The false alarms in the 1431 and 1624 data are due to a patch of grass that does not fit the background statistics. The color difference in this area is barely visible in true color images (Figure 2.8). This color variation could be due to patches of dying grass or some effect of the sun angle as these collections are later in the day than the others.



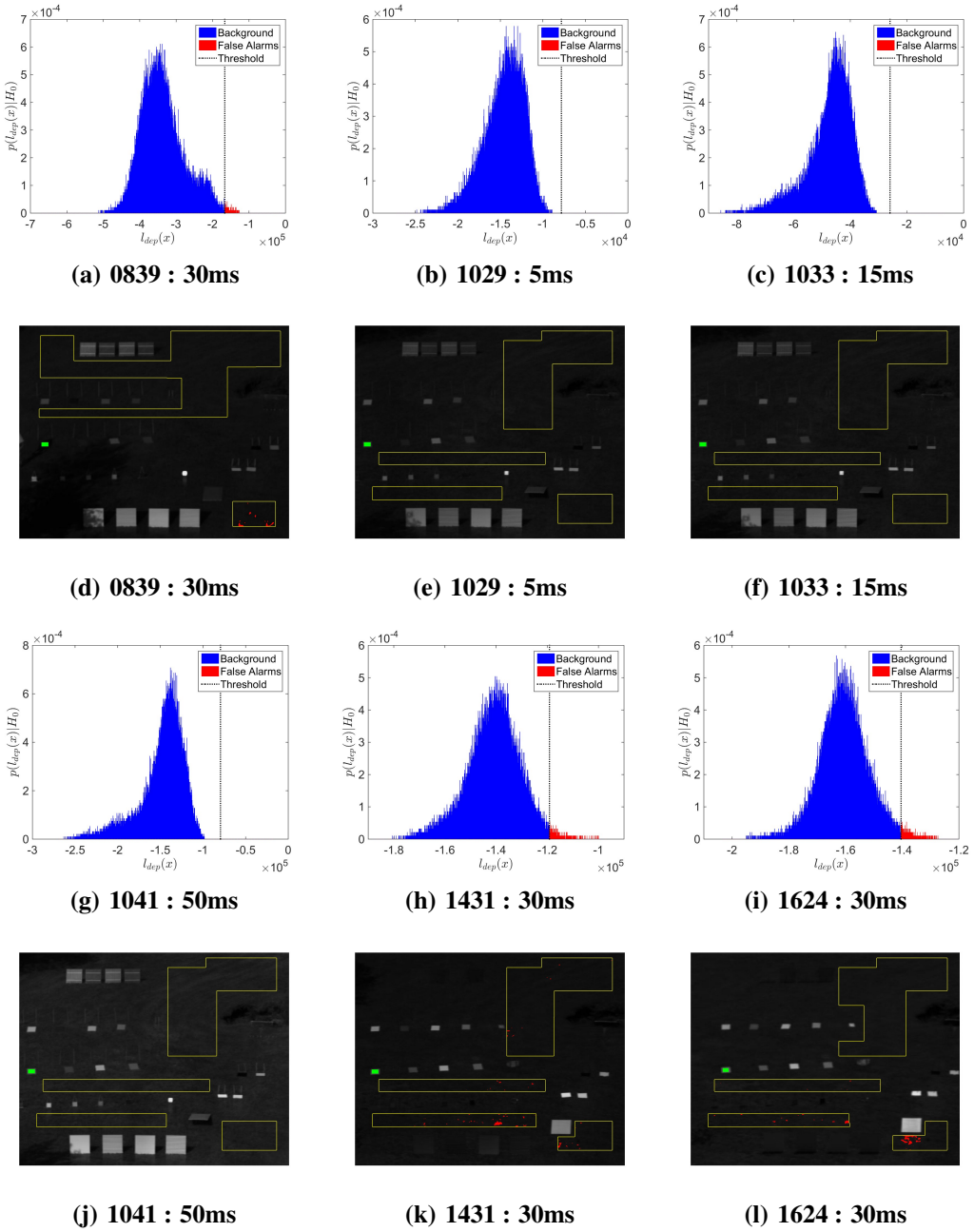
(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure 5.19: Experimental Distribution Results : Yellow, Signal-Dependent, Static. Detection statistic histograms show changes in Normality when real background data is used. Right sided skew is introduced to the 1033-1624 distributions. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure 5.20: Experimental Distribution Results : Green, Signal-Dependent, Static. Detection statistic histograms show changes in Normality when real background data is used. The left-handed skew in the 1000 results has been exaggerated by the real data. Results are reflective of a goal CFAR of 10^{-3}.

Once again recalling Section 5.2, the green panel simulation results display a strong left handed tail for the 0839-1041 data sets. The experimental results show that the inclusion of the real background exaggerates this tail in the 1000 collections while the 0839 data is now right skewed and is nearly multimodal. The only changes between the simulation results and the experimental is the true distribution of the background class. Therefore, these results suggest that the background distribution can play a significant role in the distribution of the detection statistic for some targets. In these results this is more evident in the green panel distributions than the yellow panel results.

Figure 5.21 shows the mean signature of the yellow and green panels, along with the mean background signature. The green signature displays intensity peaks near 500 and 850 nm. As

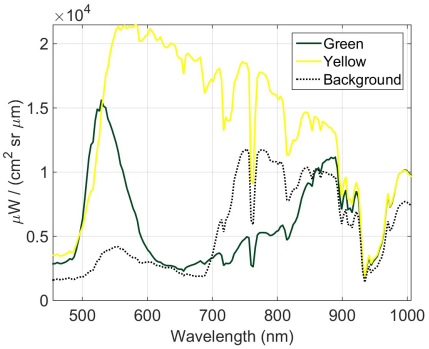
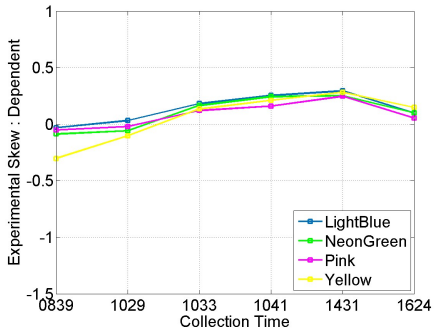


Figure 5.21: Comparison of Yellow and Green Panel Radiance. The difference in the spectral shape of the green and yellow panels leads to the differences in detection statistic distribution characteristics. The green panel two local maxima – near 550 and 850 nm – drive its distribution characteristics while the yellow panel's higher values result in a distribution whose characteristics are averaged over this region.

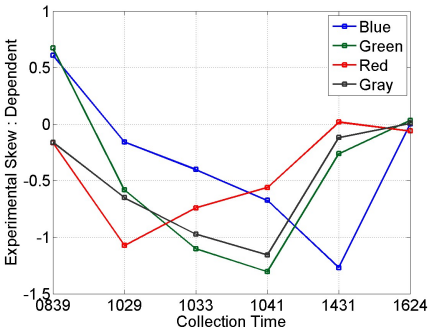
a result, these two spectral ranges are weighted more heavily in the calculation of the detection statistic. Recalling Figure 5.18, the background skew characteristics near 550 and 850 for the 0839 datacube show a heavily left-handed tail. This is clearly evident at 824nm in Figure 5.17 – very close to the second intensity peak in the green panel's signature. The reason the detection statistic distribution (Figure 5.20(a)) looks so similar to the representative background distribution at 824nm

(Figure 5.17) for this example is that the target under test heavily weighs this spectral region. This observation can be expanded to the 1029, 1033, and 1041 detection statistic distribution results. Each detection statistic skew is negatively correlated – due to the change of signs in the detection statistic – with the skew of the corresponding background in the 850nm spectral region. The yellow panel does not mimic these results because its spectral signature is higher for the entire spectral range. The detection statistics results for the yellow panel are the result of averaged distribution characteristics – which naturally lead to a more Normal result – rather than representative of a small portion like in the case of the green panel.

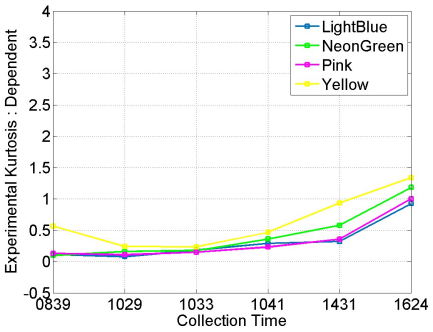
Figure 5.22 attempts to visualize the Normality trends for all of the data. The panels are split into the same two groups from before, based off of the behavior of the simulated results. These results indicate that the Group 1 panels – those whose simulation results display the highest degrees of Normality – also exhibit the experimental distributions with the best Normality measures. In general, the use of the experimental background increases the magnitude of the skew and kurtosis, thereby decreasing the Normality of the results. This is evident in the yellow panel results where the real background data has made the Normal approximation less accurate than the simulated result, but still applicable.



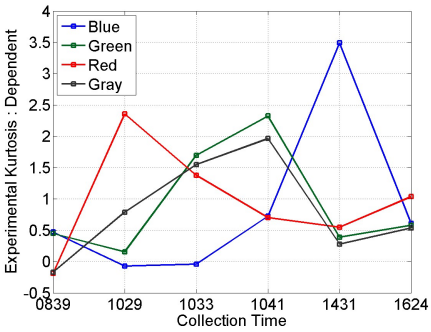
(a) Group 1 Panels : Skew



(b) Group 2 Panels : Skew



(c) Group 1 Panels : Kurtosis



(d) Group 2 Panels : Kurtosis

Figure 5.22: Experimental Normality Results : All Panels. The Group 1 panels remain highly predictable with near-Normal distributions. The real background data makes the Group 2 panels even less predictable than the simulated results. These results are representative of the signal-dependent detection statistic distribution and are nearly identical for the signal-independent distributions. See Appendix C for tables detailing the measures of Normality for each signal model.

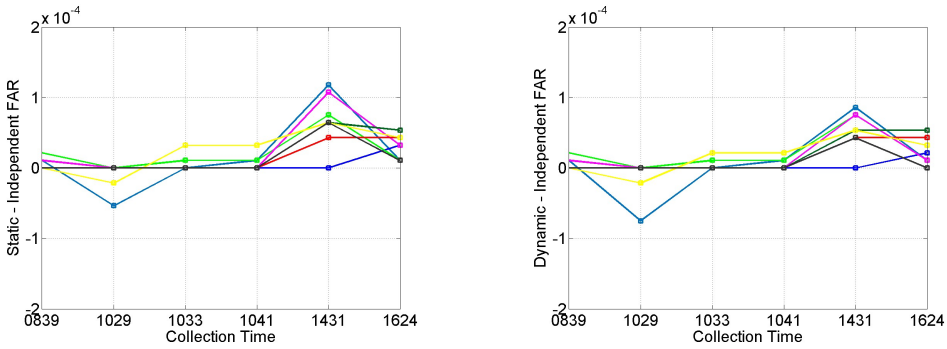
5.3.3 False Alarm Performance

The experimental results are similar the those seen in Section 5.2.3 in that the high degree of non-Normality makes comparing FAR as a function of datacube and target difficult. Instead this section focuses on finding the changes in FAR performance as a function of the noise model and thresholding method. Table 5.3 investigates the FAR of the yellow and green panels for all three thresholding methods using the real background data. Similar to the measures of Normality in Sections 5.2.2 and 5.3.2, these results suggest that the the noise model and/or thresholding method have little effect on the FAR.

Table 5.3: Experimental FAR Comparison : Yellow and Green Panels. FAR of the yellow and green panel detection statistic distributions for the 0839, 1033, and 1624 datacubes. All three thresholding methods are represented. No consistent performance improvement is noted from any of the methods. Thresholds designed for a desired CFAR of 10^{-3}.

		Signal-Independent	Static	Dynamic
0839	**Yellow**	0.91198e-03	0.91198e-03	0.91198e-03
	Green	2.57500e-03	2.57500e-03	2.57500e-03
1029	**Yellow**	1.37333e-03	1.35187e-03	1.35187e-03
	Green	0.00000e-03	0.00000e-03	0.00000e-03
1624	**Yellow**	15.40706e-03	15.44998e-03	15.43925e-03
	Green	6.48041e-03	6.53405e-03	6.53405e-03

Figure 5.23 shows the changes in FAR due to the signal-dependent detection statistic using the same methodology implemented in Section 5.2.3. Not only do these results show that the changes due to the noise model are generally minimal, when the change is large it is actually due to decreased performance of the signal-dependent model. This can be seen in the 1431 results where seven of the eight panels are greater than zero meaning the signal-independent statistic performs better. While we were able to separate the panels into the two groups based off of performance for their Normality characteristics, these groupings do not carry over into the FAR analysis.



(a) Static FAR Difference **(b) Dynamic FAR Difference**

Figure 5.23: Effect of Signal-Dependent Noise on FAR of Experimental Detection Statistics. Plots reflect the performance changes due to the signal-dependent noise thresholding techniques using the signal-independent results as reference. Note that since a small FAR is desired, negative values indicate improved performance by the signal-dependent model. Results shown are for designed CFAR of 10^{-3}

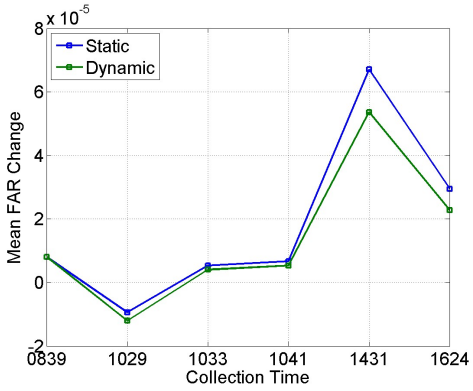


Figure 5.24: Experimental FAR : Average Change. The average FAR change attempts to represent how each data set performs overall. Results shown are for designed CFAR of 10^{-3}

The FAR changes due to the noise model do not appear to be correlated with the Normality measures of the result. By showing the average FAR change for both signal-dependent thresholding methods and each datacube the plots in Figure 5.24 attempt to communicate which of the two thresholding methods achieve overall better performance for a specific datacube. Results show that the 1029 datacube is the only cube that averages a benefit from the signal-dependent noise model.

Here, there is no trend in the 1000 datacubes to suggest what effect the SNR has on the results. Overall, the low magnitude of the performance changes due to signal-dependent noise are minimal for the data set in question. These results suggest that the noise model plays a minimal role in target detection performance for this dataset.

CHAPTER VI

CONCLUSIONS

This report has outlined the background information pertaining to HSI system modeling, target detection theory, and additive noise modeling. It then reported on the results of several experiments designed to highlight the performance of an optical system model and investigate effects of signal-dependent noise modeling. This chapter contains the conclusions drawn from these experiments and poses ideas for the future direction of this work.

6.1 Results Summary

As part of an overall effort to predict the target detection performance of a given hyperspectral sensor, this report presents the application of a system model that is used to predict the raw data output of the sensor as a function of its optical elements and collection settings. The system transfer function in (5.2) shows the ability to model the HSI system's readout with an average error of 5.5% for the entire data set. The accuracy of the model does not appear to be dependent upon the target spectra. While the model does produce impressive results, it is important to note that these results are achieved using a tuning parameter whose physical behavior is not fully understood. Results suggest that the parameter is partially dependent upon exposure time which indicates that the implemented model may overlook some natural phenomena. Additionally, the consistent spectral

shape of the parameter suggests that it may also be due to compounded error between all of the system's optical elements.

The application of the signal-independent and signal-dependent detection statistics has shown that the distribution characteristics are target dependent. The assumption is made that these distributions will be Normally distributed. This assumption in only valid for half of the panels in this dataset and needs to be addressed in future work. The non-Normal results often display an observable tail on one side of the distribution. In some cases, the non-Normal distributions result in impressively low FARs. While this is initially thought to be a positive sign for the signal-dependent detection statistic, it is quickly pointed out that distribution shapes are not dependent upon the noise model used in the detection statistic. Therefore the signal-independent results display similar FARs. Distribution characteristics such as mean, variance, skew, and kurtosis operate independently from the implemented noise model. However, minute changes in the tails of the distribution due to noise model could lead to target detection performance changes.

A second part of hyperspectral target detection performance prediction is the modeling of natural variations in an observed signal due to noise. The system transfer function provides a deterministic signal, but real signals will exhibit variations due to noise. In order to accurately predict target detection performance not only must the optical system model be accurate, but the signal model must be accurate as well. Real life signal noise displays signal-dependent characteristics. In order to truly understand the problem of target detection performance prediction, one must understand the effects of this signal-dependent noise on the target detection performance. This report compares the performance of detection statistics using both signal-independent and signal-dependent noise models in an effort to determine the effects of modeling the more computationally demanding, but accurate, signal-dependent noise and observes little performance change in terms of pD and FAR

due to the implemented noise model. *For this dataset, target detection performance has been found to be independent of the applied noise model.*

While the noise model does not seem to affect the results, it is observed that the background plays a significant role. This is especially apparent in the green panel's experimental results in Figure 5.20. This figure shows that slight changes in the sampling of the background pixels can manifest themselves as large distribution changes. This is important to the target detection performance application because it shows that without accurate background models, performance prediction suffers greatly.

6.2 Proposed Future Research

There is still a lot of room for research in modeling the HSI system. Results in this report use a flat cosine roll off function due to the targets of interest residing along the middle of the FPA. For this dataset, the simplification is accurate, but many applications may not be so restricted. Spatial dependency will need to be added to the system model functionality as well as a way to accurately model the roll off as a function of the system optics.

In addition to incorporating spatial variability into the model, the source of the tuning parameter should be further investigated. Experimental results show that the parameter changes slightly between the high and low exposure time datacubes (Figure 5.3). This could simply be natural variation of the parameter or a sign that the parameter is dependent upon exposure time. Collecting a series of integrating sphere datacubes and varying only the exposure time between collections should provide the data necessary to investigate the effects of exposure time on the parameter.

In addition to the system modeling future work, signal-dependent noise modeling research has more room to grow. The effects of signal-dependent noise on a background with less in-class variation should be investigated in order to support the conclusion that the signal-dependent noise effects

on target detection performance are minimal. The background used for these results displays a high level of natural variation. As a result – even at low SNR levels – the background variance dominates the variance due to noise. A more uniform background class would result in a lower level of in-class variance, which would allow the noise to play a larger role in the total variance of the noisy observations. In this case it is logical to conclude that the signal-dependent algorithms would have a greater effect on the target detection performance. Whether these effects would be beneficial or detrimental to overall performance remains to be seen.

Eventually, this work should be implemented together as a part of a full HSI target detection performance model. Ideally, this model would be given specifications on the reflectance of a chosen target and background. Atmospheric modeling software (not addressed in this research) would be used to model the sensor's observed pupil-plane radiance using the downwelling radiance and reflectance of the target/background. The proposed system model would then be used to model the raw data result. This would be representative of the mean spectra of the target/background. The proposed signal-dependent noise model would be used to artificially add realistic noise to this raw data. This final result could be implemented into a target detection algorithm and the performance evaluated. By implementing each of these components together into a full HSI target detection performance model it would be possible to predict how well a specific HSI system could detect a specific target in a given background and how this performance changes as a function of system components.

BIBLIOGRAPHY

[1] D. Manolakis, "Detection algorithms for hyperspectral imaging applications: a signal processing perspective," in *2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, Oct 2003, pp. 378–384.

[2] J. Kerekes and J. Baum, "Hyperspectral imaging system modeling," *Lincoln Laboratory Journal*, 2003.

[3] R. W. Basedow, D. C. Carmer, and M. E. Anderson, "HYDICE system: implementation and performance," pp. 258–267, 1995.

[4] R. W. Basedow, W. S. Aldrich, J. E. Colwell, and W. D. Kinder, "HYDICE system performance: an update," pp. 76–84, 1996.

[5] J. Pearlman, P. Barry, C. Segal, J. Shepanski, D. Beiso, and S. Carman, "Hyperion, a space-based imaging spectrometer," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 6, pp. 1160–1173, June 2003.

[6] M. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Local signal-dependent noise variance estimation from hyperspectral textural images," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 3, pp. 469–486, June 2011.

[7] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical poissonian-gaussian noise modeling and fitting for single-image raw-data," *Image Processing, IEEE Transactions on*, vol. 17, no. 10, pp. 1737–1754, Oct 2008.

[8] J. Meola, M. Eismann, R. Moses, and J. Ash, "Modeling and estimation of signal-dependent noise in hyperspectral imagery," *Applied Optics*, 2011.

[9] M. T. Eismann, *Hyperspectral Remote Sensing*. Bellingham, WA: SPIE, 2012.

[10] D. Manolakis, C. Siracusa, and G. Shaw, "Adaptive matched subspace detectors for hyperspectral imaging applications," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, vol. 5. IEEE, 2001, pp. 3153–3156.

[11] N. Acito, M. Diani, and G. Corsini, "Signal-dependent noise modeling and model parameter estimation in hyperspectral images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 49, no. 8, pp. 2957–2971, Aug 2011.

[12] *Hyperspec VNIR Datasheet*, Fitchburg, MA, 2014.

[13] *Hyperspec VS Spectrograph Certificate of Compliance*, Fitchburg, MA, 2013.

[14] *Xenoplan 1.9/35 Datasheet*, Bad Kreuznach, Germany, 2013.

[15] *Typical Spectra of Oriel Instruments: Spectral Calibration Lamps*, Stratford, CT. [Online]. Available: http://www.oriel.com

[16] *Labsphere Calibration Certificate*, North Sutton, NH, 2014.

[17] *pco.edge scientific CMOS camera*, Romulus, MI.

[18] R. P. Sperline, A. K. Knight, C. A. Gresham, D. W. Koppenaal, G. M. Hieftje, and M. B. Denton, "Read-noise characterization of focal plane array detectors via mean-variance analysis," *Appl. Spectrosc.*, vol. 59, no. 11, pp. 1315–1323, Nov 2005.

[19] G. Healey and R. V. Kondepudy, "Modeling and calibrating ccd cameras for illumination-insensitive machine vision," pp. 121–132, 1992. [Online]. Available: http://dx.doi.org/10.1117/12.57974

[20] H. Stark and J. W.Woods, "Some important random processes," in *Probability and Random Processes with Applications to Signal Processing*, 3rd ed. Upper Saddle River, NJ: Princeton Hall, 2002, ch. 7.2.

[21] L.-R. Gao, B. Zhang, X. Zhang, W.-J. Zhang, and Q.-X. Tong, "A new operational method for estimating noise in hyperspectral images," *Geoscience and Remote Sensing Letters, IEEE*, vol. 5, no. 1, pp. 83–87, Jan 2008.

[22] M. L. Uss, B. Vozel, V. V. Lukin, and K. Chehdi, "Maximum likelihood estimation of spatially correlated signal-dependent noise in hyperspectral images," *Optical Engineering*, vol. 51, no. 11, pp. 111 712–1–111 712–11, 2012. [Online]. Available: http://dx.doi.org/10.1117/1.OE.51.11.111712

[23] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, Aug 1992.

[24] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, Sept 1962.

[25] S. M. Kay, "Statistical Decision Theory I," in *Fundamentals of Statistical Signal Processing Detection Theory*. Upper Saddle River, NJ: Prentice-Hall, Inc, 1998, ch. 3.

[26] K. Fukunaga, *Statistical Pattern Recognition*. San Diego, CA: Academic Press, 1990.

[27] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer Science & Business Media, 1994.

[28] X. Yu and I. Reed, "Comparative performance analysis of adaptive multiband detectors," *IEEE Trans. Signal Proc.*, vol. 41, no. 8, pp. 2639–2656, 1993.

[29] K. B. Peterson and M. S. Pedersen, "The matrix cookbook," November 2012.

[30] A. M. Mathai and S. B. Provost, "Moment generating functions and characteristic functions," in *Quadratic Forms in Random Variables*. New York, NY: Marcel Dekker, Inc, 1992, ch. 3.2.

APPENDIX A

DERIVATIONS OF THEORETICAL MEAN AND VARIANCE OF DETECTION STATISTICS

A.1 Null Hypothesis : $\mathrm{x} \sim \mathcal{N}(\boldsymbol{\mu}_b, \boldsymbol{\Sigma}_0)$

The derivation of (4.30) is detailed below, beginning with a restatement of (4.28),

$$\begin{aligned}\sigma_0^2 = E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_0\Big] - 4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big]\ldots \\ + 4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_0\Big] + 2W_0E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big]\ldots \\ + 4E\Big[(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})^2\Big|H_0\Big] - 8E\Big[(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_0\Big] - 4W_0\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_0\Big]\ldots \\ + 4E\Big[(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})^2\Big|H_0\Big] + 4W_0E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_0\Big] + W_0^2. \quad \text{(A.1.1)}\end{aligned}$$

Each expectation in (A.1.1) is calculated individually below. Beginning with the first term, we invoke (4.20) to find,

$$\begin{aligned}E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_0\Big] = \ldots \\ 2Tr\big((\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_0(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_0\big) + 4\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_0(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b\big)\ldots \\ + \Big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b + Tr\big((\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_0\big)\Big)^2. \quad \text{(A.1.2)}\end{aligned}$$

The covariance matrices are distributed and simplified resulting in,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_0\Big] = 2Tr\big((I-\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2\big)+4\big(\boldsymbol{\mu}_b^T(I-\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b\big)\ldots$$
$$+\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b+K-Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\big)^2. \quad \text{(A.1.3)}$$

The trace operation is expanded and simplified, the inner product of the second term is carried out, and the final squared term is calculated to find,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_0\Big] = 2K+2Tr\big((\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2\big)-4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots$$
$$-8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b\big)^2+2K(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b)\ldots$$
$$-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b\big)+Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2-2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)+K^2 \quad \text{(A.1.4)}$$

Finally, squared terms are carried out and other terms are simplified to contain a single covariance matrix when possible,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_0\Big] = 2K+2Tr\big((\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2\big)-4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots$$
$$-8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)^2-2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big)\ldots$$
$$+\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big)^2+2K\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)-2K\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big)-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)\ldots$$
$$+2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big)+Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2-2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)+K^2. \quad \text{(A.1.5)}$$

The next expectation is first broken into two terms given by,

$$-4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big] = -4E\Big[(\mathbf{x}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big]\ldots$$
$$+4E\Big[(\mathbf{x}^T\boldsymbol{\Sigma}_n^{-1}\mathbf{x})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big]. \quad \text{(A.1.6)}$$

The expectations are calculated according to (4.20) once again,

$$-4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big]=\dots$$
$$-8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\dots$$
$$+8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b). \quad \text{(A.1.7)}$$

The result is simplified to acheive the final result given by,

$$-4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_0\Big]=-8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2\dots$$
$$+8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b). \quad \text{(A.1.8)}$$

The next term is calculated in manner analogous to (A.1.8) resulting in,

$$4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_0\Big]=\dots$$
$$8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\dots$$
$$-8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)-4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) \quad \text{(A.1.9)}$$

The fourth term is separated into a difference of expectations before (4.17) is applied and W_0 is expanded,

$$2W_0E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big]=2W_0\big(E[\mathbf{x}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\big|H_0]-E[\mathbf{x}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\big|H_0]\big), \quad \text{(A.1.10)}$$

$$2W_0E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big]=2\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b-2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)-K\big)\times\dots$$
$$\big(Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_0)+\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b-Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)-\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big). \quad \text{(A.1.11)}$$

The constant is distributed and trace function simplified resulting in,

$$2W_0E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big] = \dots$$
$$\big(2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)-2K\big)\times\dots$$
$$\big(K+\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b-Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)-\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big). \quad \text{(A.1.12)}$$

The two terms are distributed to find,

$$2W_0E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big] = 2K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+2K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)-4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\dots$$
$$+2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)-2K^2+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\dots$$
$$-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-2K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\dots$$
$$-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0^{-1})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\dots$$
$$-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2+2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)-2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)-2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)^2\dots$$
$$+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+2K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b). \quad \text{(A.1.13)}$$

Cancellation of terms simplifies to a final result given by,

$$2W_0E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_0\Big] = 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)-4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+4KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\dots$$
$$-2K^2+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\dots$$
$$-4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)+4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2\dots$$
$$-2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)^2+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b). \quad \text{(A.1.14)}$$

The fifth expectation is found by applying (4.18) and results in,

$$4E\Big[\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\big)^2\Big|H_0\Big] = 4\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)+4\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)^2. \quad \text{(A.1.15)}$$

Similarly, (4.18) is applied to the sixth expectation resulting in the following,

$$-8E\Big[\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\big)\big(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\big)\Big|H_0\Big] = -8\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\big) - 8\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)\big(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big). \quad \text{(A.1.16)}$$

By substituting W_0 and calculating the expectation, the seventh term can be given by,

$$\begin{aligned}-4W_0E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_0\Big] &= \dots \\ &- 4\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b - 2\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s} + Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - K\big)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big). \quad \text{(A.1.17)}\end{aligned}$$

Distribution of the coefficient and separating the covariances yields the following,

$$\begin{aligned}-4W_0E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_0\Big] &= (\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\times\dots \\ &\big(-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) + 4K\big). \quad \text{(A.1.18)}\end{aligned}$$

The last term is distributed leading to a final result of

$$\begin{aligned}-4W_0E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_0\Big] &= -4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2 - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) + \dots \\ &8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) + 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b). \quad \text{(A.1.19)}\end{aligned}$$

The eighth expectation is calculated by applying (4.18) to find,

$$4E\Big[\big(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\big)^2\Big|H_0\Big] = 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)^2. \quad \text{(A.1.20)}$$

By substituting W_0 and calculating the expectation, the ninth term is found to be,

$$4W_0E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_0\Big] = \dots$$
$$4\big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b - 2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - K\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b). \quad \text{(A.1.21)}$$

After distributing the coefficient and separating the covariances the result is given by,

$$4W_0E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_0\Big] = \dots$$
$$\big(4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - 4K\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) \quad \text{(A.1.22)}$$

The last term is distributed leading to a final result of,

$$4W_0E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_0\Big] = 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)\dots$$
$$- 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 + 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) \quad \text{(A.1.23)}$$

The final term does not involve the calculation of an expectation. We expand W_0,

$$W_0^2 = \big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b - 2\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s} + Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - K\big)^2 \quad \text{(A.1.24)}$$

carry out the square,

$$W_0^2 = \big(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b\big)^2\dots$$
$$- 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b) + 2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b)\dots$$
$$- 2K(\boldsymbol{\mu}_b^T(\boldsymbol{\Sigma}_0^{-1}+\boldsymbol{\Sigma}_1^{-1})\boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\dots$$
$$+ Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2 - 2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) + K^2, \quad \text{(A.1.25)}$$

and multiply out individual terms in order to find the final result given by,

$$W_0^2 = (\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_b)^2 + 2(\boldsymbol{\mu}_b \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_b)(\boldsymbol{\mu}_b \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b) + (\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b)^2 \dots$$
$$- 4(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \mathbf{s})(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_b) - 4(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \mathbf{s})(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b) + 2Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_b) \dots$$
$$+ 2Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b) - 2K(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_b) - 2K(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \mathbf{s})^2 \dots$$
$$- 4Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \mathbf{s}) + 4K(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \mathbf{s}) + Tr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0)^2 - 2KTr(\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\Sigma}_0) \dots$$
$$+ K^2. \quad \text{(A.1.26)}$$

Substitution of the calculated expectations (A.1.5), (A.1.8), (A.1.9), (A.1.15), (A.1.16), (A.1.19),

(A.1.20), (A.1.23), and (A.1.26) into (A.1.1) results in,

$$\begin{aligned}
\sigma_0^2 = {} & 2K + 2Tr\big((\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2\big) - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + \big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)^2 - 2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big) + \big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big)^2\ldots \\
& + 2K\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big) - 2K\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big) - 2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)\ldots \\
& + 2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big) + Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2 - 2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) + K^2 - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots \\
& - 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2 + 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots \\
& - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) - 2K^2 + 2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2\ldots \\
& - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\ldots \\
& - 2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2 - 2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)^2 + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2 - 8\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\big) - 8\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b\big)\big(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b\big) - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2\ldots \\
& - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) + 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)^2 + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 + 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)\ldots \\
& + (\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)^2 + 2(\boldsymbol{\mu}_b\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)(\boldsymbol{\mu}_b\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + (\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b)^2\ldots \\
& - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots \\
& + 2Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) - 2K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 2K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2\ldots \\
& - 4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0)^2 - 2KTr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0) + K^2
\end{aligned} \tag{A.1.27}$$

The result from combining and/or canceling terms is found to be,

which can be rewritten into the final result,

A.2 Alternate Hypothesis : $\mathbf{x} \sim \mathcal{N}(\mathbf{s}, \boldsymbol{\Sigma}_1)$

The derivation of (4.31) is detailed below. The LRT (4.11) is rewritten as follows,

$$l(\mathbf{x}) = \mathbf{x}^T \boldsymbol{\Sigma}_0^{-1} \mathbf{x} - 2(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \mathbf{x}) + \boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\mu}_b - (\mathbf{x} - \mathbf{s})^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}) \tag{A.2.1}$$

We define,

$$\mu_1 = E[l(\mathbf{x}|H_1)] \tag{A.2.2}$$

and recall that $\mathbf{x} \sim \mathcal{N}(\mathbf{s}, \boldsymbol{\Sigma}_1)$ under H_1. According to (4.17) the first term's expected value follows,

$$E\big[\mathbf{x}^T \boldsymbol{\Sigma}_0^{-1} \mathbf{x}\big| H_1\big] = Tr(\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_1) + \mathbf{s}^T \boldsymbol{\Sigma}_0^{-1} \mathbf{s}. \tag{A.2.3}$$

The following two terms have expected values given by,

$$E\big[-2(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \mathbf{x}) + \boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b | H_1\big] = -2(\boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_0^{-1} \mathbf{s}) + \boldsymbol{\mu}_b^T \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_b. \tag{A.2.4}$$

The final term results in an expected Chi-Squared distribution [29],

$$(\mathbf{x} - \mathbf{s})^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}) \sim \chi^2_K \tag{A.2.5}$$

which has an expected value of K. These results are summed to find that

$$\mu_1 = (\mathbf{s} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{s} - \boldsymbol{\mu}_b) + Tr(\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_1) - K. \tag{A.2.6}$$

Next, the variance is calculated according to,

$$\sigma_1^2 = E\Big[\big((\mathbf{x} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) - (\mathbf{x} - \mathbf{s})^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}) - \mu_1\big)^2 \Big| H_1\Big]. \tag{A.2.7}$$

When μ_1 is substituted back in the result is given by,

$$\sigma_1^2 = E\Big[\big((\mathbf{x} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{x} - \boldsymbol{\mu}_b) - (\mathbf{x} - \mathbf{s})^T \boldsymbol{\Sigma}_1^{-1} (\mathbf{x} - \mathbf{s}) + (\mathbf{s} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_0^{-1} (\mathbf{s} - \boldsymbol{\mu}_b) + Tr(\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_1) - K\big)^2 \Big| H_1\Big]. \tag{A.2.8}$$

Individual terms are then distributed resulting in,

$$\sigma_1^2 = E\Big[\big(\mathbf{x}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x} - 2\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x} + \boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b - \mathbf{x}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x} + 2\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x} - \mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\ldots$$
$$- \mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s} + 2\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s} - \boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b - Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + K\big)^2\Big|H_1\Big] \quad \text{(A.2.9)}$$

which is further simplified to,

$$\sigma_1^2 = E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x} - 2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}) + 2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}) + W_1\big)^2\Big|H_1\Big]. \quad \text{(A.2.10)}$$

Where W_1 is made up of all terms not dependent upon $\mathbf{x}$ and is defined as follows,

$$W_1 = -\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_0^{-1})\mathbf{s} + 2\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s} - Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + K. \quad \text{(A.2.11)}$$

The square is distributed accordingly,

$$\sigma_1^2 = E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2 - 4\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}) + 4\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\ldots$$
$$+ 2W_1\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})^2 - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\ldots$$
$$- 4W_1(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})^2 + 4W_1(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}) + W_1^2\Big|H_1\Big] \quad \text{(A.2.12)}$$

followed by the distribution of the expectation,

$$\sigma_1^2 = E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_1\Big] - 4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_1\Big]\ldots$$
$$+ 4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_1\Big] + 2W_1E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_1\Big]\ldots$$
$$+ 4E\Big[\big(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\big)^2\Big|H_1\Big] - 8E\Big[(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_1\Big] - 4W_1E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_1\Big]\ldots$$
$$+ 4E\Big[(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})^2\Big|H_1\Big] + 4W_1E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_1\Big] + W_1^2. \quad \text{(A.2.13)}$$

Each expectation in (A.2.13) is calculated individually in the following pages. Beginning with the first term, we invoke (4.20) to find,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_1\Big] = \ldots$$
$$2Tr\big((\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_1\big) + 4\big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{s}\big)\ldots$$
$$+ \Big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\mathbf{s} + Tr\big((\boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_1^{-1})\boldsymbol{\Sigma}_1\big)\Big)^2. \quad \text{(A.2.14)}$$

The covariance matrices are distributed and simplified resulting in,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_1\Big] = 2Tr\big((\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1-I)^2\big)\ldots$$
$$+4\big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1-I)(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{s}\big)+\big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{s}+Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)-K\big)^2. \quad \text{(A.2.15)}$$

The trace operation is expanded and simplified, the inner product of the second term is carried out, and the final squared term is calculated to find,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_1\Big] = 2Tr\big((\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2\big)-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+2K+4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$
$$-8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+\big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{s}\big)^2+2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)\big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{s}\big)\ldots$$
$$-2K\big(\mathbf{s}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{s}\big)+Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2-2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+K^2. \quad \text{(A.2.16)}$$

Finally, squared terms are carried out and other terms are simplified to contain a single covariance matrix when possible,

$$E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)^2\Big|H_1\Big] = 2Tr\big((\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2\big)-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+2K+4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1})\mathbf{s}\ldots$$
$$-8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2-2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2\ldots$$
$$+2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-2K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$
$$+2K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2-2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+K^2. \quad \text{(A.2.17)}$$

The next expectation is first broken into two terms given by,

$$-4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_1\Big] = -4E\Big[(\mathbf{x}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_1\Big]\ldots$$
$$+4E\Big[(\mathbf{x}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_1\Big] \quad \text{(A.2.18)}$$

The expectations are calculated according to (4.19) once again,

$$-4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_1\Big] = -8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots$$
$$-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+8(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots$$
$$+4Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}). \quad \text{(A.2.19)}$$

The result is simplified to achieve the final result given by,

$$-4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})\Big|H_1\Big] = 8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)-8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots$$
$$+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$
$$+4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}). \quad \text{(A.2.20)}$$

The next term is calculated in a manner analogous to (A.2.20) resulting in,

$$4E\Big[\big(\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_1\Big] = 8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-8(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\ldots$$
$$+4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2. \quad \text{(A.2.21)}$$

The fourth term is separated into a difference of expectations before (4.17) is applied and W_1 is expanded,

$$2W_1E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_1\Big] = 2W_1\big(E[\mathbf{x}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}|H_1]-E[\mathbf{x}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}|H_1]\big), \quad \text{(A.2.22)}$$

$$2W_1E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_1\Big] = \ldots$$
$$2\left(-\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s}+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+K\right)\times\ldots$$
$$\big(Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}-Tr(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_1)-\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\big). \quad \text{(A.2.23)}$$

The constant is distributed and trace function simplified resulting in,

$$2W_1E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_1\Big] = \dots$$
$$\big(-2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+2K\big)\times\dots$$
$$\big(Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}-K-\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}\big). \quad \text{(A.2.24)}$$

The two terms are distributed to find,

$$2W_1E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_1\Big] = -2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots$$
$$+2K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2-2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2+2K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots$$
$$+2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots$$
$$-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2-2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)\dots$$
$$+2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+2K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots$$
$$-2K^2-2K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}). \quad \text{(A.2.25)}$$

Cancellation of terms simplifies to a final result given by,

$$2W_1E\Big[\mathbf{x}^T(\boldsymbol{\Sigma}_0^{-1}-\boldsymbol{\Sigma}_1^{-1})\mathbf{x}\Big|H_1\Big] = 2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2\dots$$
$$+4K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots$$
$$-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})-4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2+4KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)\dots$$
$$-2K^2. \quad \text{(A.2.26)}$$

The fifth expectation is found by applying (4.18) and results in,

$$4E\Big[(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})^2\Big|H_1\Big] = 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2. \quad \text{(A.2.27)}$$

Similarly, (4.18) is applied to the sixth expectation resulting in the following,

$$-8E\Big[(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})\Big|H_1\Big] = -8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}). \tag{A.2.28}$$

By substituting W_1 and calculating the expectation, the seventh term can be given by,

$$\begin{aligned}-4W_1E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_1\Big] &= \dots \\ &- 4\left(-\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s} + 2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + K\right)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}).\end{aligned} \tag{A.2.29}$$

Distribution of the coefficient and separating the covariances yields the following,

$$\begin{aligned}-4W_1E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_1\Big] &= 4\big(\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s}\big)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots \\ &+ 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\end{aligned} \tag{A.2.30}$$

The last term is distributed leading to a final result of

$$\begin{aligned}-4W_1E\Big[\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{x}\Big|H_1\Big] &= 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\dots \\ &- 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 + 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}).\end{aligned} \tag{A.2.31}$$

The eighth expectation is calculated by applying (4.18) to find,

$$4E\Big[(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x})^2\Big|H_1\Big] = 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2. \tag{A.2.32}$$

By substituting W_1 and calculating the expectation, the ninth term is found to be,

$$4W_1E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_1\Big]\dots$$
$$= 4\big(-\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s}+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+K\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}). \quad \text{(A.2.33)}$$

After distributing the coefficient and separating the covariances the result is given by,

$$4W_1E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_1\Big] = -4\big(\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s}\big)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\dots$$
$$-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}). \quad \text{(A.2.34)}$$

The last term is distributed leading to a final result of,

$$4W_1E\Big[\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{x}\Big|H_1\Big] = -4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2-4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\dots$$
$$-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})+4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}). \quad \text{(A.2.35)}$$

The final term does not involve the calculation of an expectation. We expand W_1,

$$W_1^2 = \big(-\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s}+2(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})-Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+K\big)^2 \quad \text{(A.2.36)}$$

and carry out the square,

$$W_1^2 = \big(\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s}\big)^2-4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s})\dots$$
$$+2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s})-2K(\mathbf{s}^T(\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_0^{-1})\mathbf{s})+4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2\dots$$
$$-4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})+Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2\dots$$
$$-2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)+K^2, \quad \text{(A.2.37)}$$

and multiply out individual terms in order to find the final result given by,

$$W_1^2 = (\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 + 2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$
$$+ 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 2K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + (\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 - 4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$
$$+ 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 2K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 - 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$
$$+ 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2 - 2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + K^2. \quad \text{(A.2.38)}$$

Substitution of the calculated expectations (A.2.17), (A.2.20), (A.2.21), (A.2.27), (A.2.28), (A.2.31), (A.2.32), (A.2.35), and (A.2.38) into (A.2.13) results in,

$$\sigma_1^2 = 2Tr\big((\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2\big) - 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + 2K + 4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1})\mathbf{s} - 8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$+ 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + (\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 - 2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + (\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2\ldots$$

$$+ 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 2K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 2K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\ldots$$

$$+ Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2 - 2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + K^2 + 8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) - 8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b)\ldots$$

$$+ 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$+ 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 8(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 8(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\ldots$$

$$+ 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 + 2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2\ldots$$

$$- 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 2(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 + 4K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$+ 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2\ldots$$

$$+ 4KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) - 2K^2 + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_b) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 - 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$- 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$- 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 + 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2\ldots$$

$$- 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 - 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 8(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})\ldots$$

$$+ 4K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + (\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})^2 + 2(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 4(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$+ 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) - 2K(\mathbf{s}^T\boldsymbol{\Sigma}_1^{-1}\mathbf{s}) + (\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 - 4(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$+ 2Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) - 2K(\mathbf{s}^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + 4(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})^2 - 4Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s})\ldots$$

$$+ 4K(\boldsymbol{\mu}_b^T\boldsymbol{\Sigma}_0^{-1}\mathbf{s}) + Tr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1)^2 - 2KTr(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_1) + K^2. \quad \text{(A.2.39)}$$

APPENDIX B

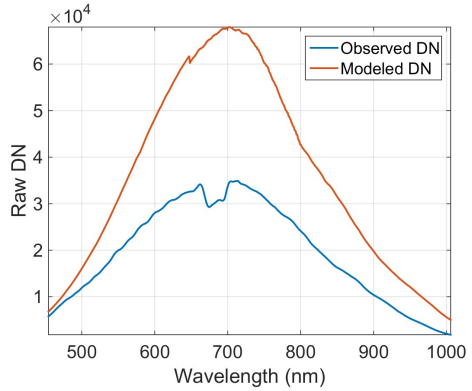
SYSTEM TRANSMISSION MODELING RESULTS - SUPPLEMENTAL FIGURES

B.1 Untuned System Transmission Modeling : Integrating Sphere Data

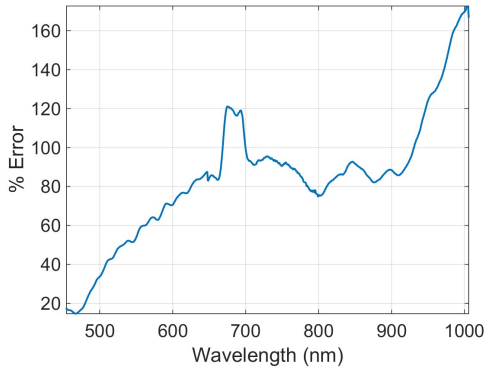
Figures B.1 and B.2 look at the average error of the system transmission model as a function of the f/# of the foreoptics. They reinforce the results in the main body of research, that the spectral shape of the error remains consistent. The error metric is computed,

$$\frac{Modeled - Observed}{Observed}. \tag{B.1.1}$$

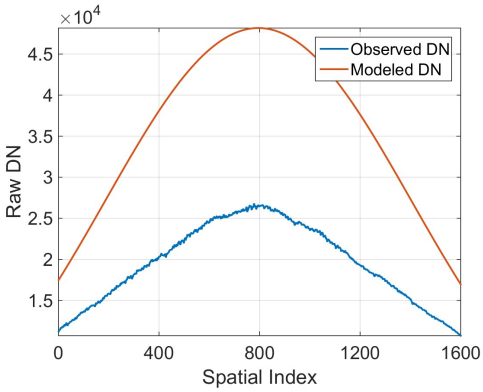
It is important to note that these results look at the entire active FPA. The results from the scene data come from a small region of interest within the FPA and the target panels are within an even smaller ROI. In this small area the cosine rolloff is able to be modeled as a flat function with a high degree of accuracy.



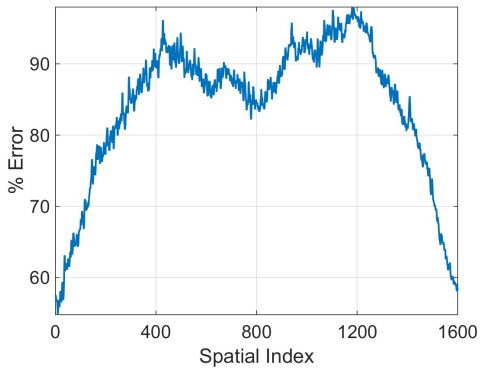
(a) Spectral Mean Comparison



(b) Spectral Mean Error

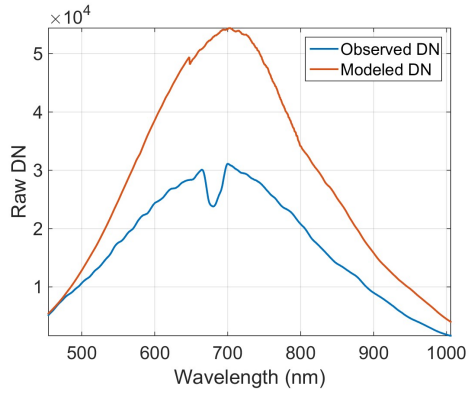


(c) Spatial Mean Comparison

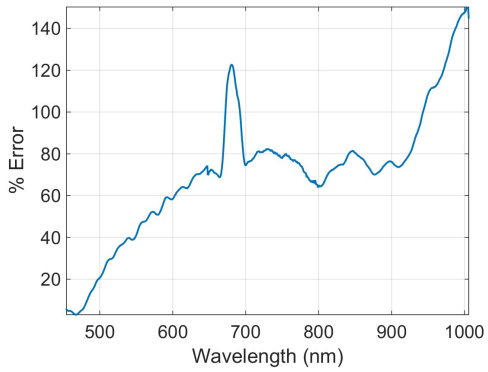


(d) Spatial Mean Error

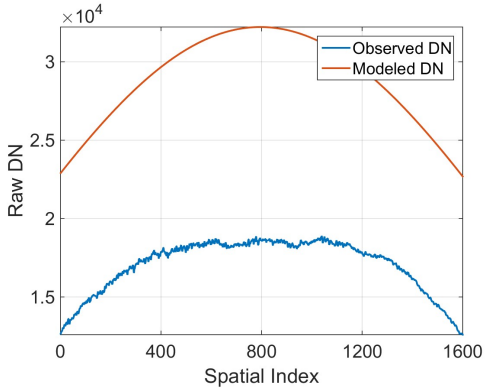
Figure B.1: Untuned System Modeling Results : 35mm Lens at f/1.9. According to the error plots the modeled data is 84% higher than the observed. The error due to cosine rolloff gets quite large at the extremes of the spatial indices.



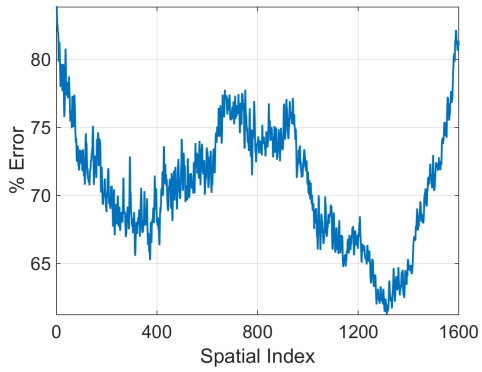
(a) Spectral Mean Comparison



(b) Spectral Mean Error



(c) Spatial Mean Comparison



(d) Spatial Mean Error

Figure B.2: Untuned System Modeling Results : 35mm Lens at f/2.8. These results come from integrating sphere data. According to the error plots the modeled data is 71% higher than the observed. The higher f/# results in a flatter cosine rolloff. While the spatial results are more accurate, there is still a 20% swing in error.

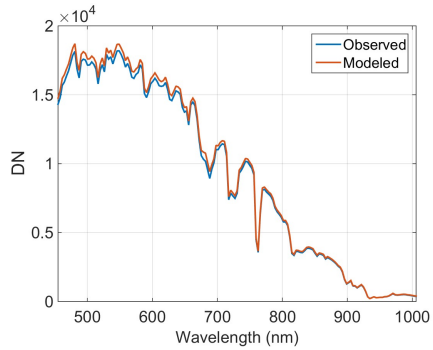
B.2 Tuned System Modeling Data : Scene

The following results are from the scene data, after the calculation of the tuning parameter. Table B.1 shows the absolute percent error between the obeserved and modeled spectra for each panel in each datacube. Overall the results show a high degree of accuracy, with the average error at 5%. Following the error results are figures comparing the observed and modeled signals represented in Table B.1.

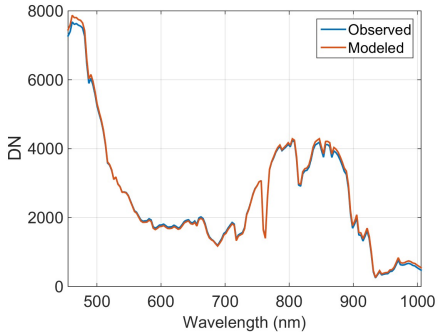
Finally, Figure B.9 expands on the results of Table B.1 by showing the spectral error – between the modeled and observed signas – of all results. These results show that our system modeling method is very accurate for all but the 1029 datacube for wavelengths less than 900nm. After 900nm all results are significantly less accurate.

Table B.1: Tuned System Model Error : All Panels. Table contains the mean absolute percent error between the observed and modeled raw data spectra. All values higher than 10% are highlighted in red. The highest observed error is 15.4%.

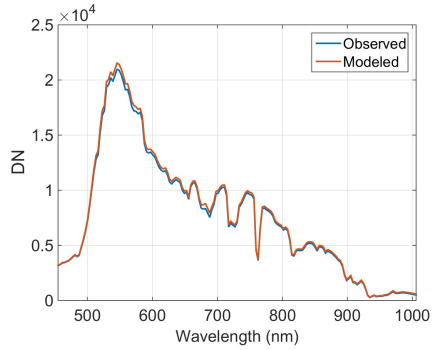
	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray	Avg.
0839	0.02632	0.03074	0.03814	0.02659	0.06310	0.02025	0.01999	0.02059	0.03072
1029	0.09013	0.15024	0.05779	0.08592	0.11958	0.11665	0.06930	0.15352	0.10539
1033	0.02104	0.03683	0.03863	0.02034	0.02504	0.02668	0.02153	0.03392	0.02800
1041	0.09259	0.09180	0.10206	0.09111	0.07711	0.07560	0.08250	0.05520	0.08350
1431	0.04407	0.04827	0.05513	0.04394	0.03889	0.03819	0.03799	0.01744	0.04049
1624	0.04098	0.04912	0.05084	0.04100	0.03940	0.03738	0.03790	0.01839	0.03938
Avg.	0.05252	0.06783	0.05710	0.05149	0.06052	0.05246	0.04487	0.04984	0.05458



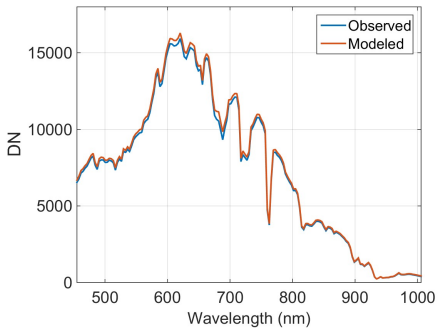
(a) Light Blue



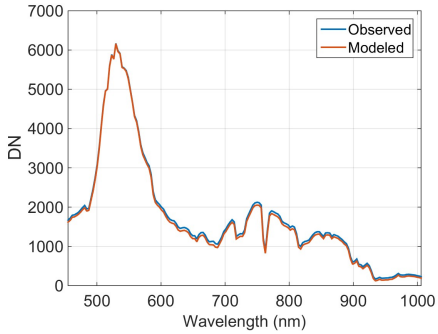
(b) Blue



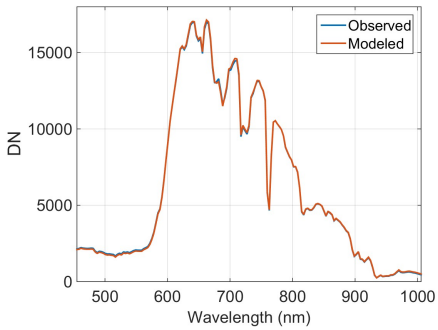
(c) Neon



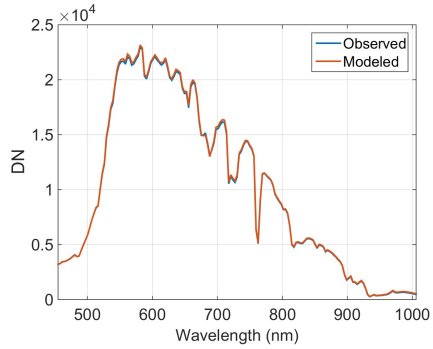
(d) Pink



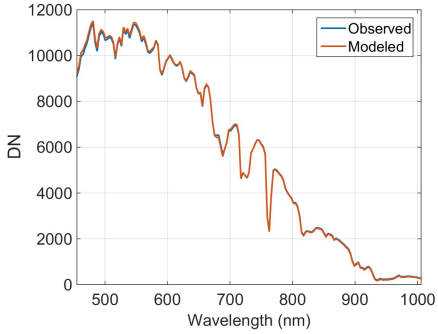
(e) Green



(f) Red

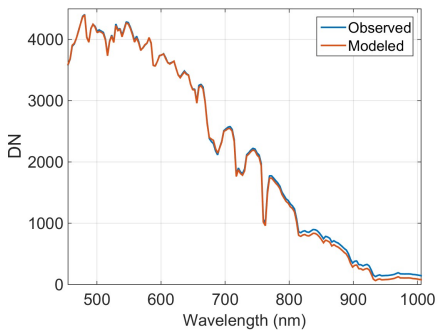


(g) Yellow

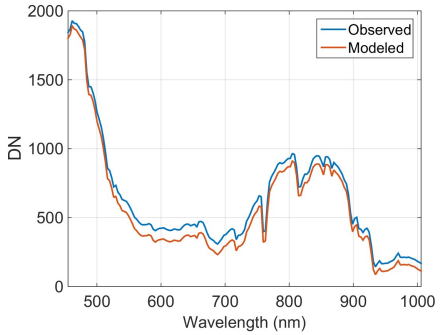


(h) Gray

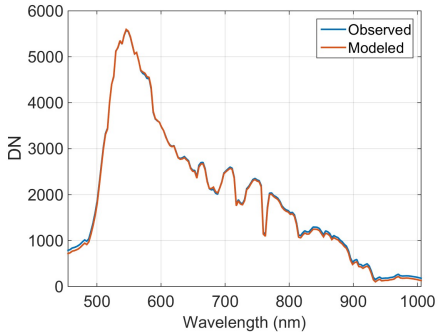
Figure B.3: Tuned System Modeling Results : 0839 Datacube.



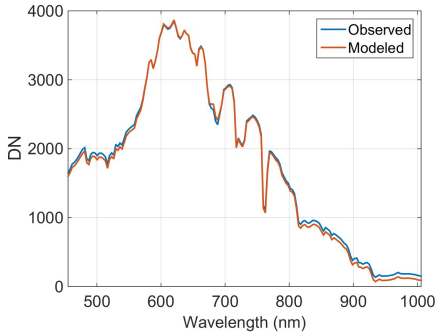
(a) Light Blue



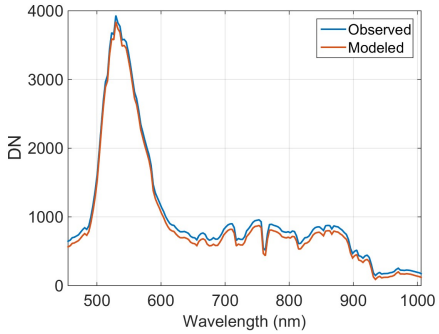
(b) Blue



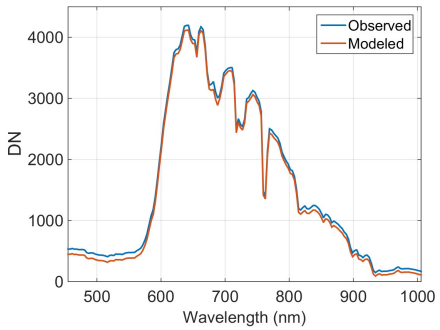
(c) Neon



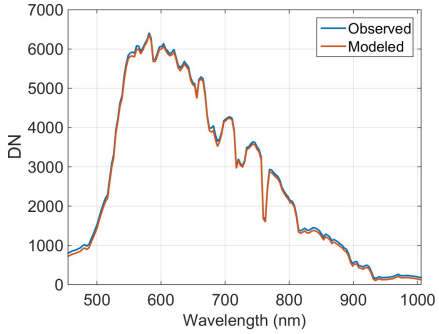
(d) Pink



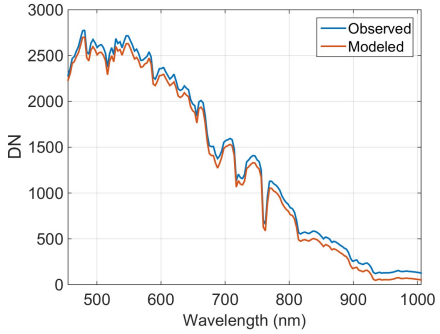
(e) Green



(f) Red

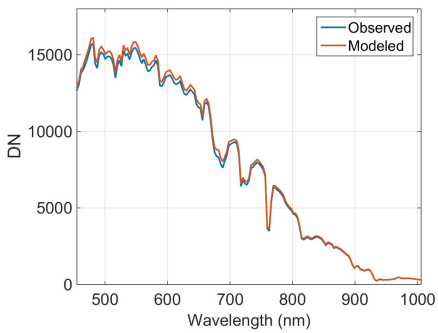


(g) Yellow

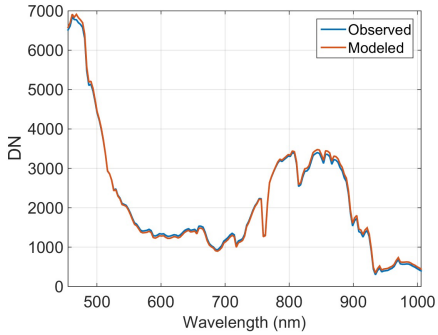


(h) Gray

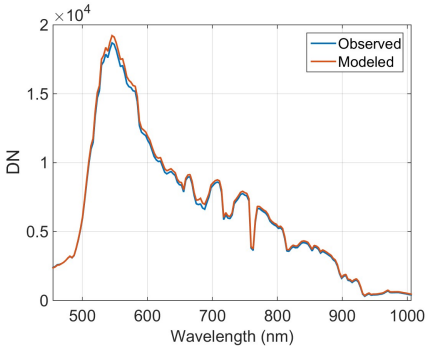
Figure B.4: Tuned System Modeling Results : 1029 Datacube.



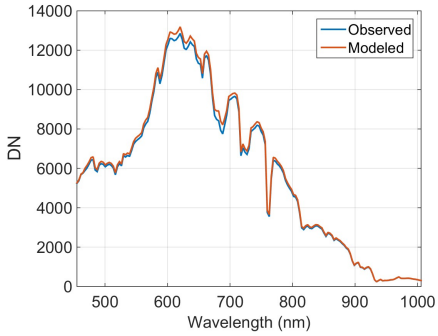
(a) Light Blue



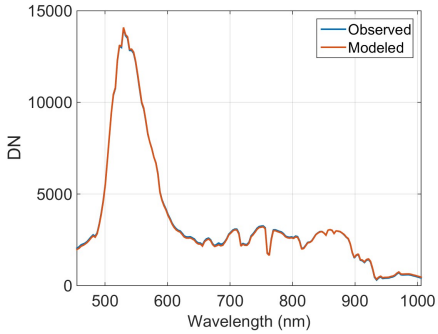
(b) Blue



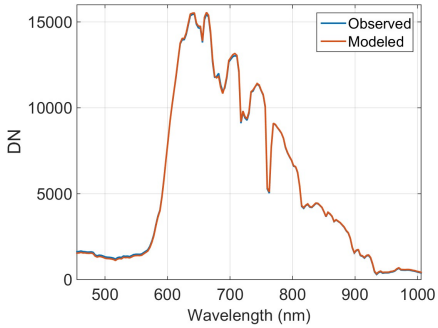
(c) Neon



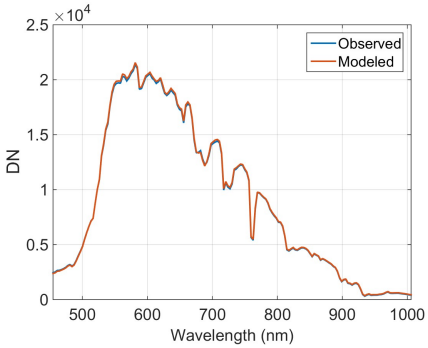
(d) Pink



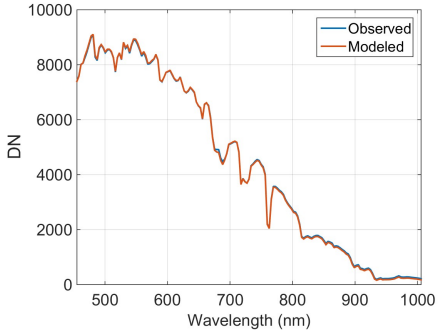
(e) Green



(f) Red

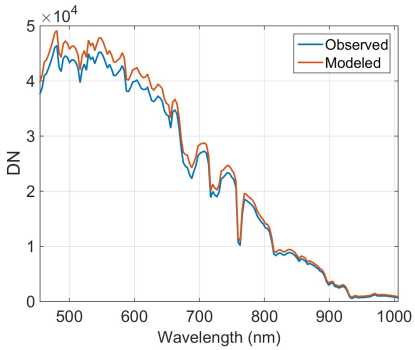


(g) Yellow

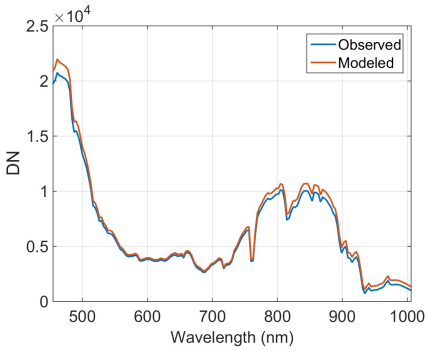


(h) Gray

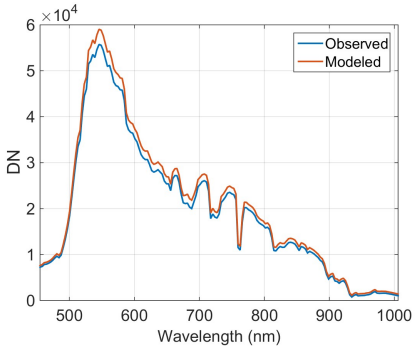
Figure B.5: Tuned System Modeling Results : 1033 Datacube.



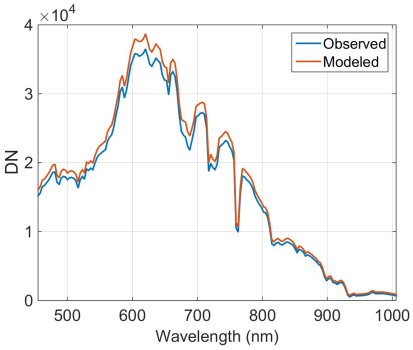
(a) Light Blue



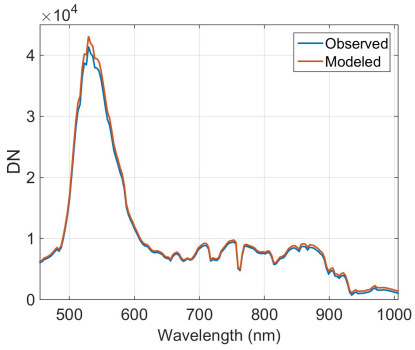
(b) Blue



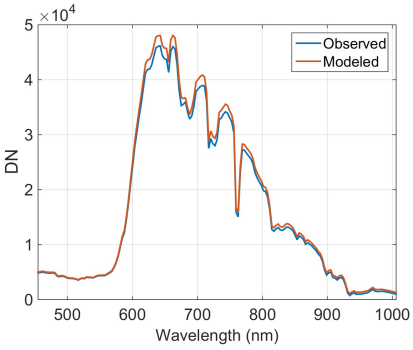
(c) Neon



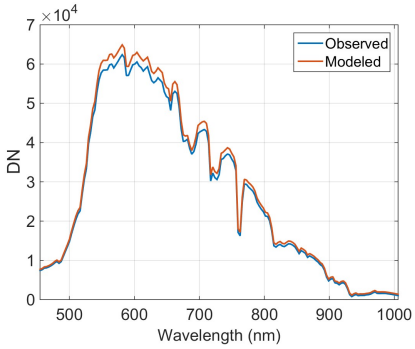
(d) Pink



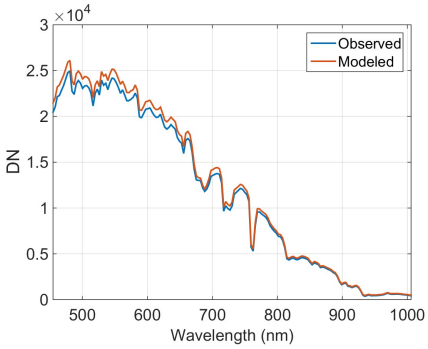
(e) Green



(f) Red

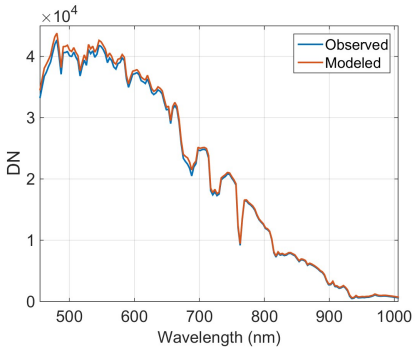


(g) Yellow

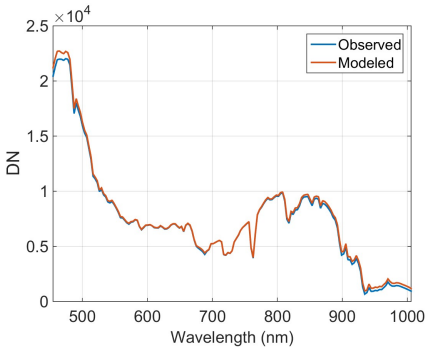


(h) Gray

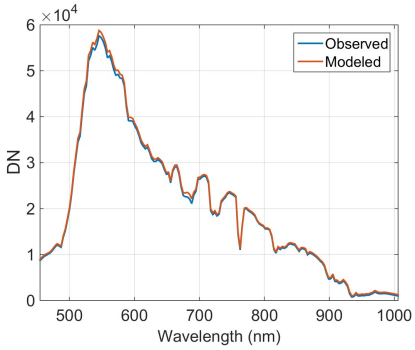
Figure B.6: Tuned System Modeling Results : 1041 Datacube.



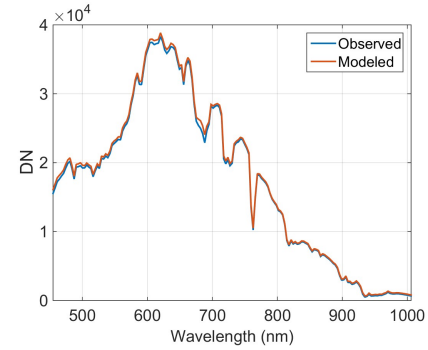
(a) Light Blue



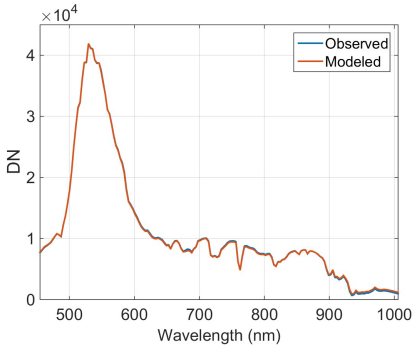
(b) Blue



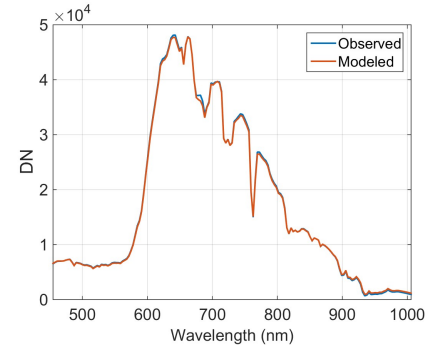
(c) Neon



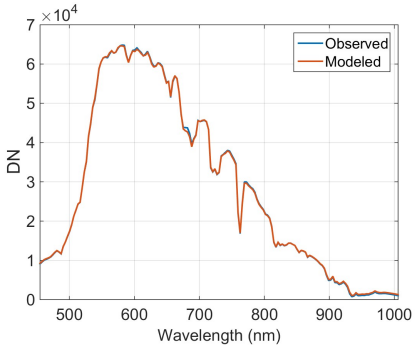
(d) Pink



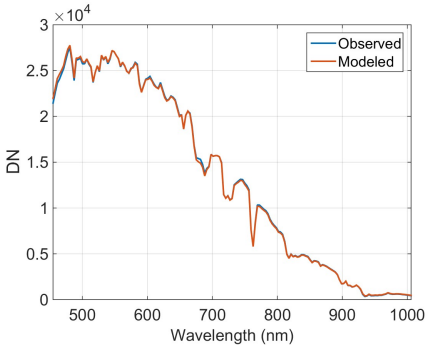
(e) Green



(f) Red

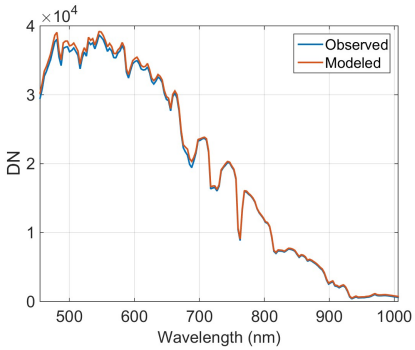


(g) Yellow

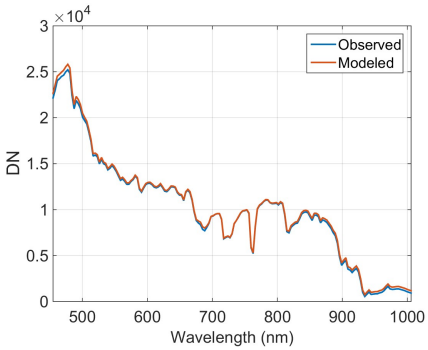


(h) Gray

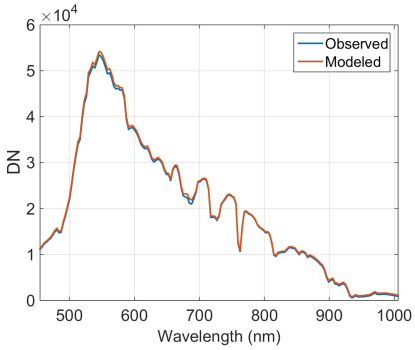
Figure B.7: Tuned System Modeling Results : 1431 Datacube.



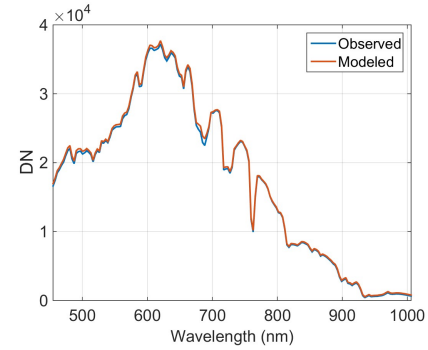
(a) Light Blue



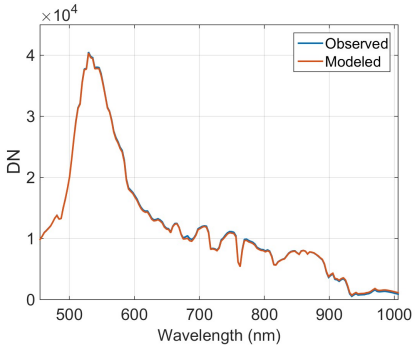
(b) Blue



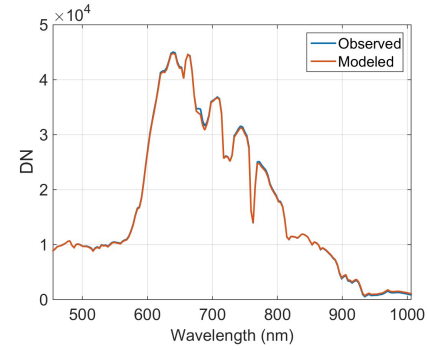
(c) Neon



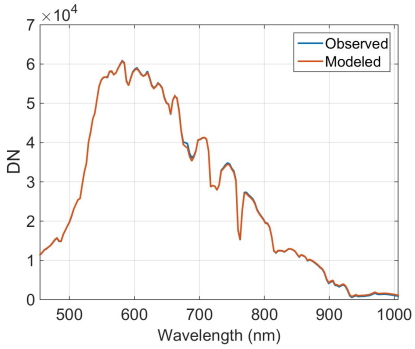
(d) Pink



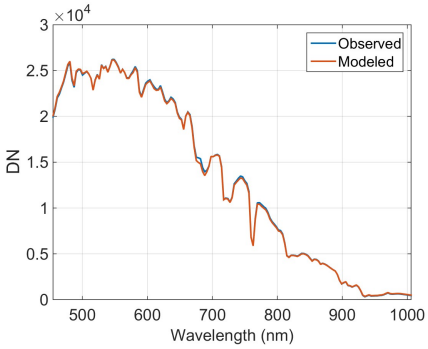
(e) Green



(f) Red

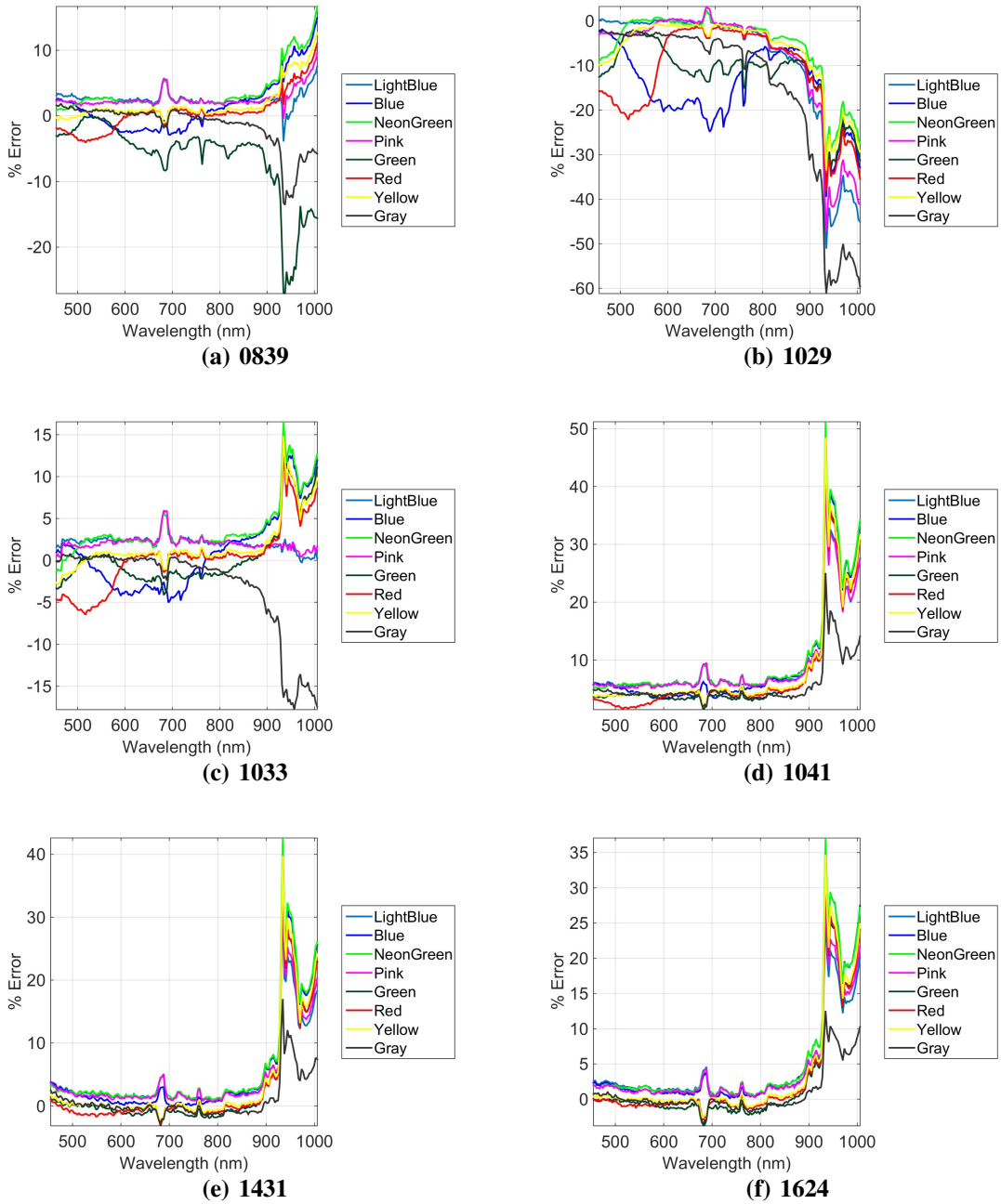


(g) Yellow



(h) Gray

Figure B.8: Tuned System Modeling Results : 1624 Datacube.



(a) 0839 (b) 1029

(c) 1033 (d) 1041

(e) 1431 (f) 1624

Figure B.9: Tuned System Modeling Results : Error Plots. These results show that for all but the 1029 datacube the results are between $\pm$ 10% for wavelengths less than 900 nm. The 1029 datacube does not fit this behavior, but that corresponds with the results in the main body of this research.

APPENDIX C

SIMULATED DATA MODELING RESULTS - SUPPLEMENTAL TABLES

Tables C.1-C.6 contain the FAR results of all of the simulated data. Each table is representative of a single datacube. The rows labeled **Indep.**, **Static**, and **Dynamic** represent the FAR for the signal-independent detection statistic, the signal-dependent static threshold, and the signal-dependent dynamic threshold respectively. The skew and kurtosis values are also included to give reference to the Normality of the resultant distribution. These values are representative of the signal-dependent distribution, but the main body of research shows that these values do not change much between the signal-independent and signal-dependent distributions.

Note that all FAR results have been multiplied by a factor of 10^3 *to save room. This means that a FAR of* 1 *indicates a perfect match with the designed CFAR of* 10^{-3}.

	Light Blue	**Blue**	**Neon**	**Pink**	**Green**	**Red**	**Yellow**	**Gray**
Indep.	0.52860	0.08380	0.26985	0.41720	0.15070	0.00020	0.44225	0.04905
Static	0.52865	0.08380	0.27000	0.41745	0.15070	0.00020	0.44265	0.04900
Dynamic	0.52845	0.08385	0.27000	0.41740	0.15075	0.00020	0.44270	0.04900
Skew	-0.11548	-0.30450	-0.21366	-0.15011	-0.25792	-0.74834	-0.15090	-0.36168
Kurtosis	0.02261	0.12586	0.06998	0.03431	0.09001	0.89054	0.04702	0.20043

Table C.1: Simulated FAR and Normality Results : 0839 Datacube. Normalized empirical FAR of simulated data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	0.50335	0.35475	0.34220	0.32540	0.00075	0.06465	0.46545	0.14160
Static	0.50325	0.35460	0.34240	0.32515	0.00075	0.06485	0.46525	0.14215
Dynamic	0.50350	0.35465	0.34230	0.32510	0.00075	0.06380	0.46520	0.14210
Skew	-0.12540	-0.21640	-0.18230	-0.18774	-0.50766	-0.72470	-0.13606	-0.29854
Kurtosis	0.02925	0.09496	0.05712	0.06336	0.37990	1.58564	0.02963	0.17116

Table C.2: Simulated FAR and Normality Results : 1029 Datacube. Normalized empirical FAR of simulated data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	0.61805	0.01495	0.43890	0.44155	0.00030	0.31460	0.32725	0.08925
Static	0.61825	0.01495	0.43900	0.44170	0.00030	0.31415	0.32685	0.08940
Dynamic	0.61830	0.01500	0.43900	0.44170	0.00030	0.31485	0.32775	0.08925
Skew	-0.09101	-0.39376	-0.14320	-0.14246	-0.50238	-0.29826	-0.28555	-0.33488
Kurtosis	0.01385	0.20693	0.03368	0.03262	0.35548	0.38111	0.35557	0.19477

Table C.3: Simulated FAR and Normality Results : 1033 Datacube. Normalized empirical FAR of simulated data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	0.62885	0.00970	0.46555	0.46350	0.15905	0.09750	0.54770	0.03700
Static	0.62875	0.00970	0.46540	0.46360	0.15905	0.09730	0.54760	0.03700
Dynamic	0.62895	0.00970	0.46540	0.46360	0.15905	0.09745	0.54760	0.03700
Skew	-0.08799	-0.40557	-0.13690	-0.13458	-0.39159	-0.51510	-0.11150	-0.39603
Kurtosis	0.01219	0.22021	0.03397	0.02590	0.26327	0.74333	0.02284	0.23552

Table C.4: Simulated FAR and Normality Results : 1041 Datacube. Normalized empirical FAR of simulated data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	0.78060	0.00100	0.70385	0.70000	0.25655	0.40780	0.75375	0.45305
Static	0.78050	0.00100	0.70400	0.69995	0.25660	0.40785	0.75385	0.45265
Dynamic	0.78040	0.00100	0.70380	0.70025	0.25670	0.40785	0.75385	0.45290
Skew	-0.04889	-0.51749	-0.07016	-0.06998	-0.23497	-0.15354	-0.05707	-0.15543
Kurtosis	0.00703	0.39743	0.01121	0.01196	0.11802	0.03425	0.00738	0.06547

Table C.5: Simulated FAR and Normality Results : 1431 Datacube. Normalized empirical FAR of simulated data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

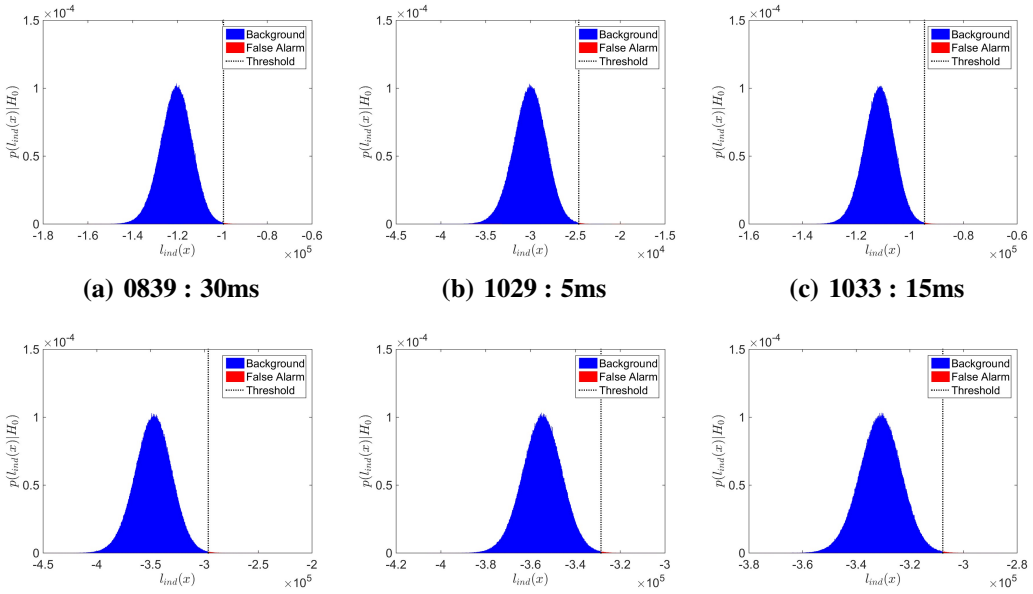
	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	0.75370	0.27530	0.64940	0.67055	0.38500	0.43700	0.65825	0.37250
Static	0.75390	0.27530	0.64945	0.67035	0.38500	0.43730	0.65825	0.37265
Dynamic	0.75380	0.27510	0.64960	0.67035	0.38500	0.43730	0.65830	0.37285
Skew	-0.05503	-0.23168	-0.08733	-0.08025	-0.18651	-0.14433	-0.07951	-0.19786
Kurtosis	0.00909	0.12909	0.02183	0.02008	0.10015	0.03517	0.01358	0.11010

Table C.6: Simulated FAR and Normality Results : 1624 Datacube. Normalized empirical FAR of simulated data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

APPENDIX D

SIMULATED DATA MODELING RESULTS - SUPPLEMENTAL FIGURES

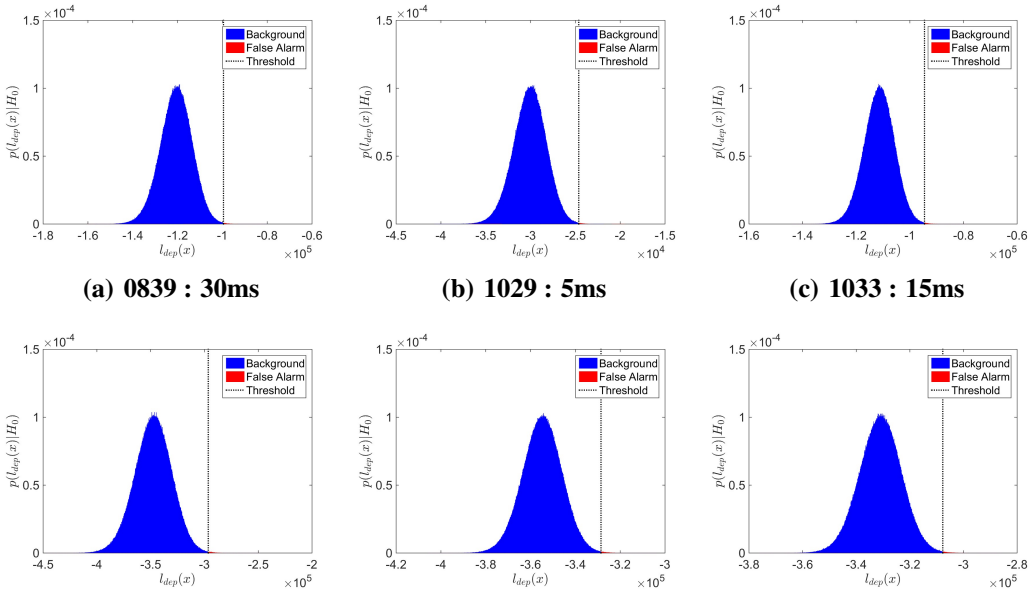
Figures D.1 - D.12 display the similarities between the signal-independent and signal-dependent detection statistics for all results. These results are calculated in an identical manner to Figures 5.6-5.9 in the main body of research.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

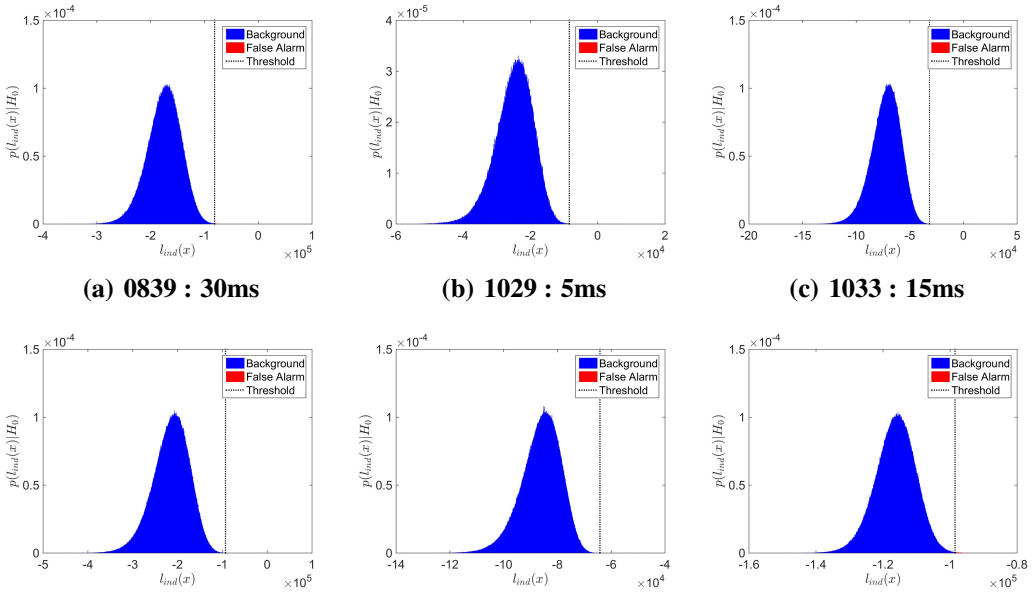
Figure D.1: Simulated Distribution Results : Light Blue, Signal-Independent. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

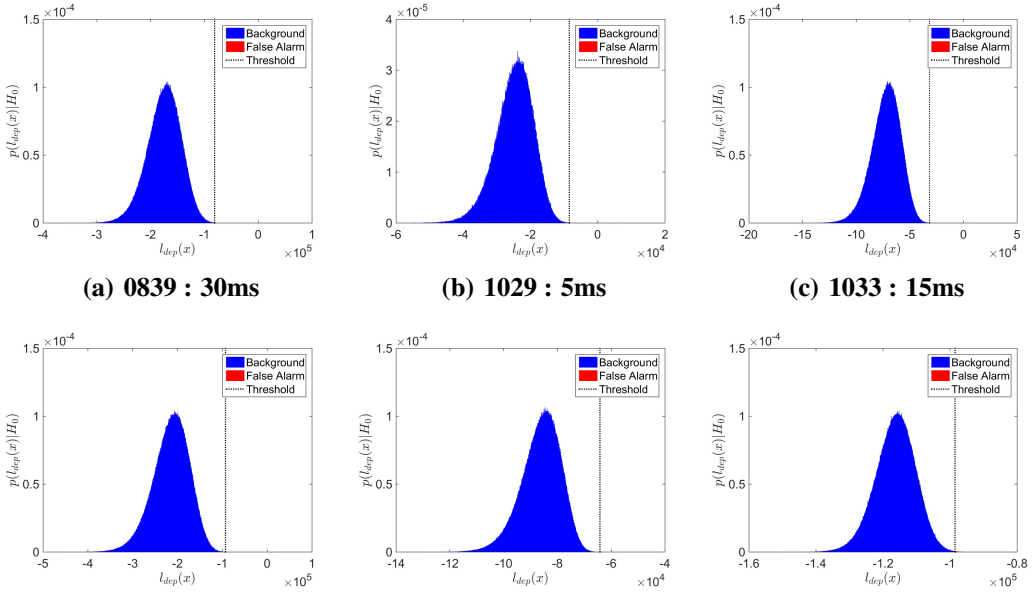
Figure D.2: Simulated Distribution Results : Light Blue, Signal-Dependent. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

(d) 1041 : 50ms (e) 1431 : 30ms (f) 1624 : 30ms

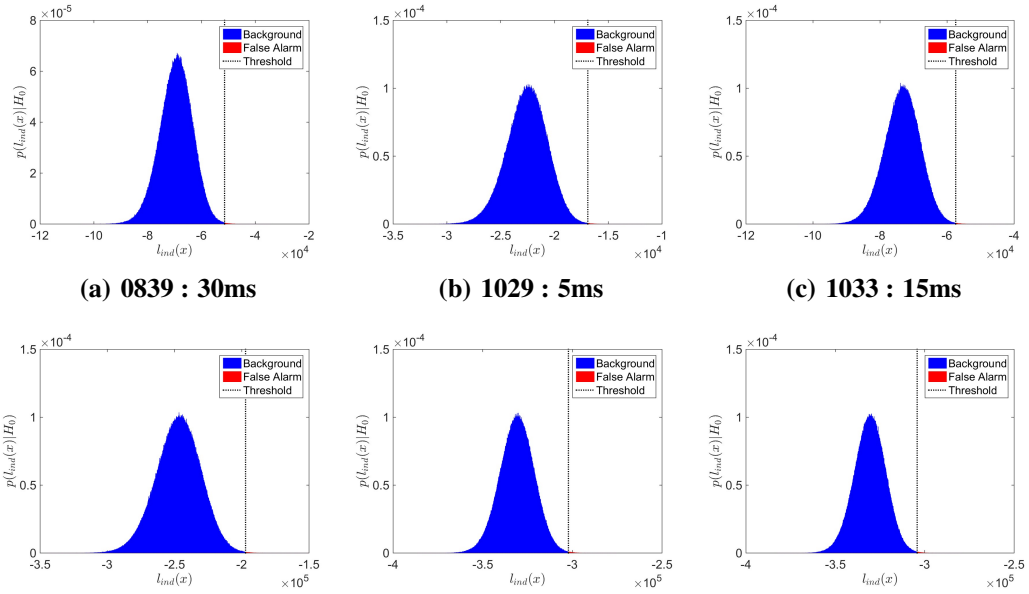
Figure D.3: Simulated Distribution Results : Blue, Signal-Independent. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

(d) 1041 : 50ms (e) 1431 : 30ms (f) 1624 : 30ms

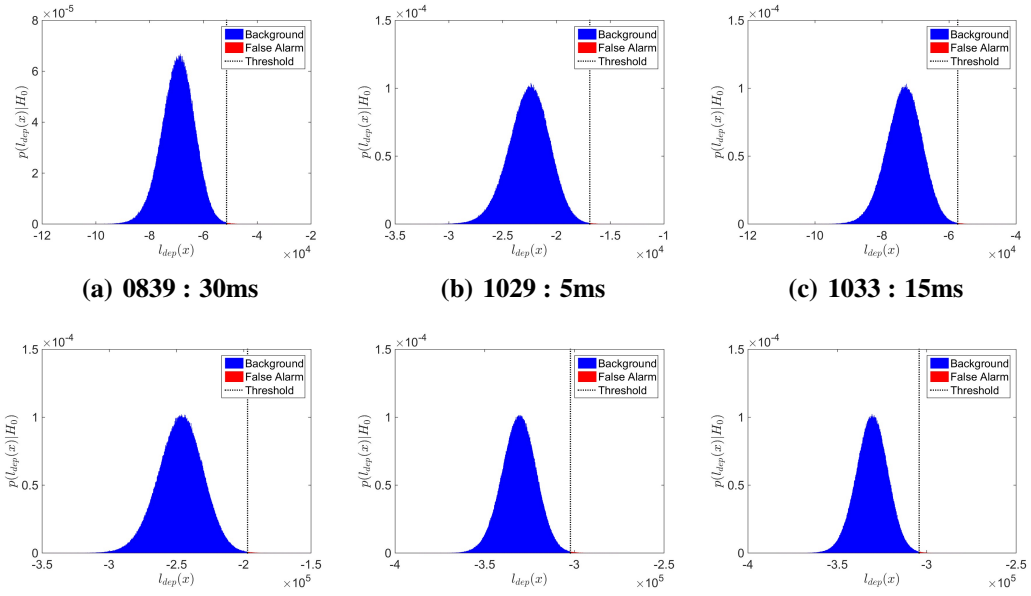
Figure D.4: Simulated Distribution Results : Blue, Signal-Dependent. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

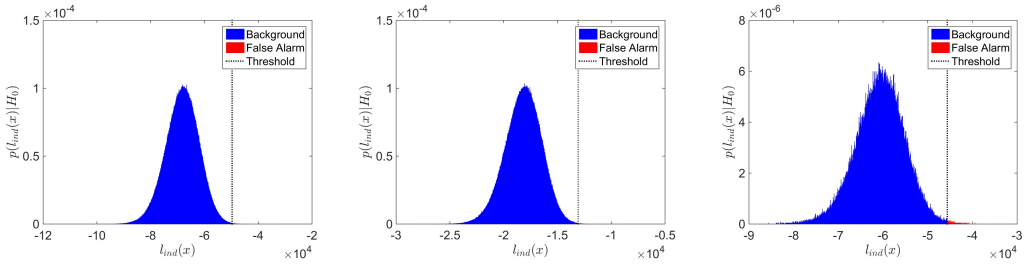
Figure D.5: Simulated Distribution Results : Neon Green, Signal-Independent. Results are reflective of a goal CFAR of 10^{-3}.



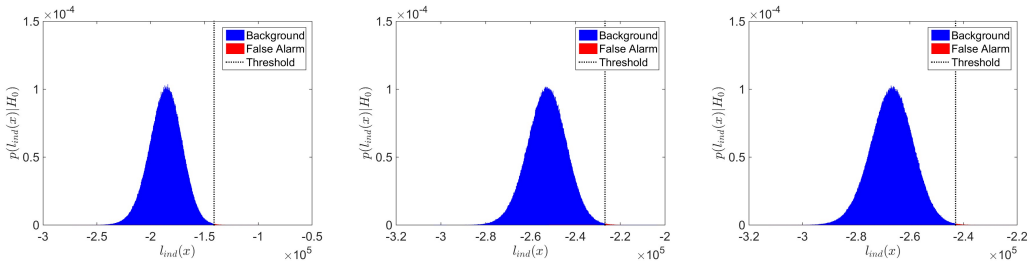
(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure D.6: Simulated Distribution Results : Neon Green, Signal-Dependent. Results are reflective of a goal CFAR of 10^{-3}.

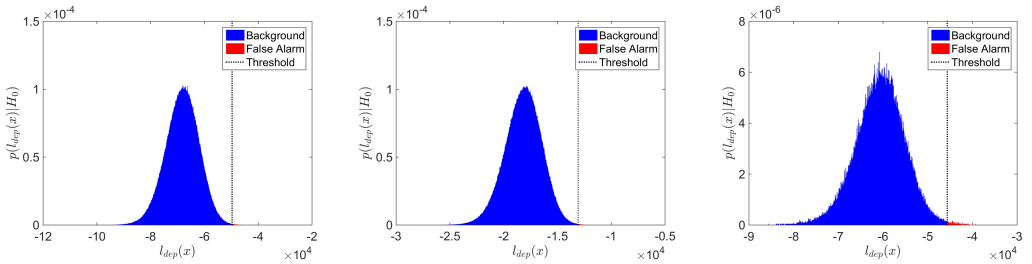


(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

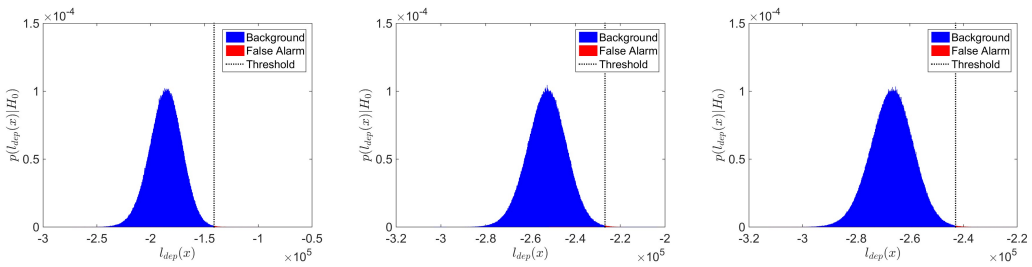


(d) 1041 : 50ms (e) 1431 : 30ms (f) 1624 : 30ms

Figure D.7: Simulated Distribution Results : Pink, Signal-Independent. Results are reflective of a goal CFAR of 10^{-3}.

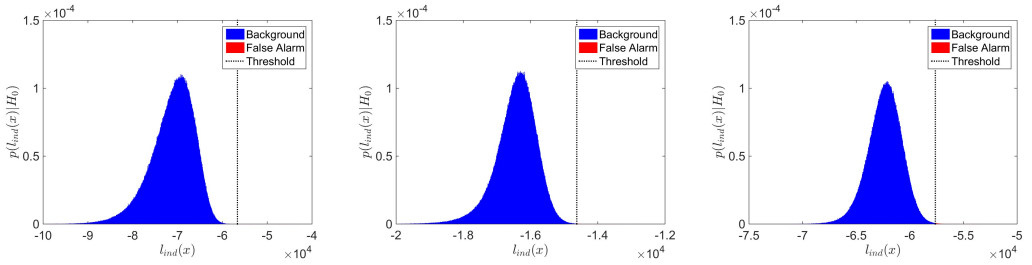


(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

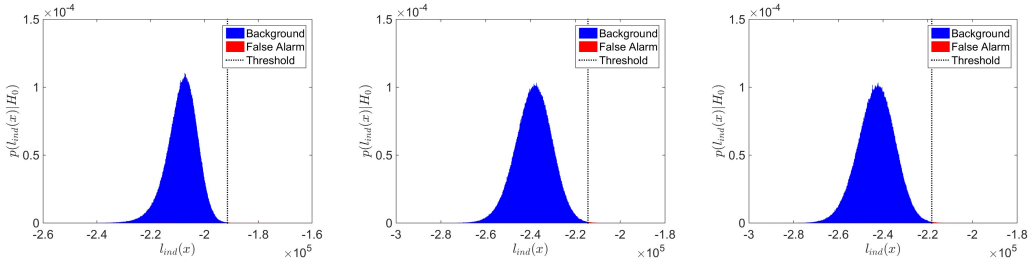


(d) 1041 : 50ms (e) 1431 : 30ms (f) 1624 : 30ms

Figure D.8: Simulated Distribution Results : Pink, Signal-Dependent. Results are reflective of a goal CFAR of 10^{-3}.

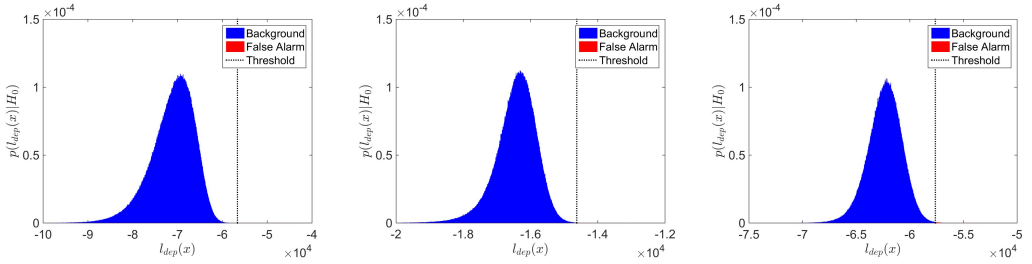


(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

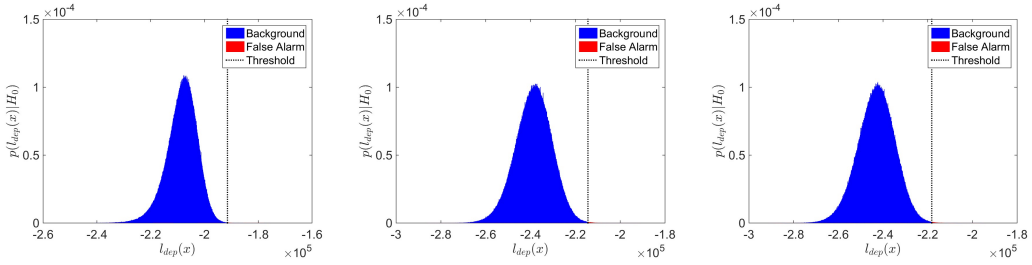


(d) 1041 : 50ms (e) 1431 : 30ms (f) 1624 : 30ms

Figure D.9: Simulated Distribution Results : Red, Signal-Independent. Results are reflective of a goal CFAR of 10^{-3}.

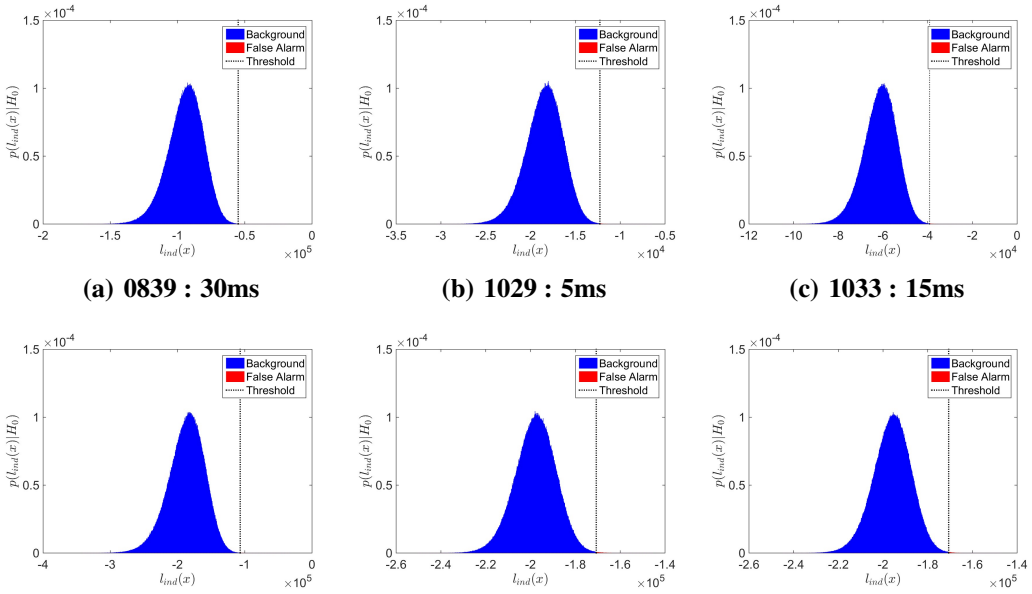


(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms



(d) 1041 : 50ms (e) 1431 : 30ms (f) 1624 : 30ms

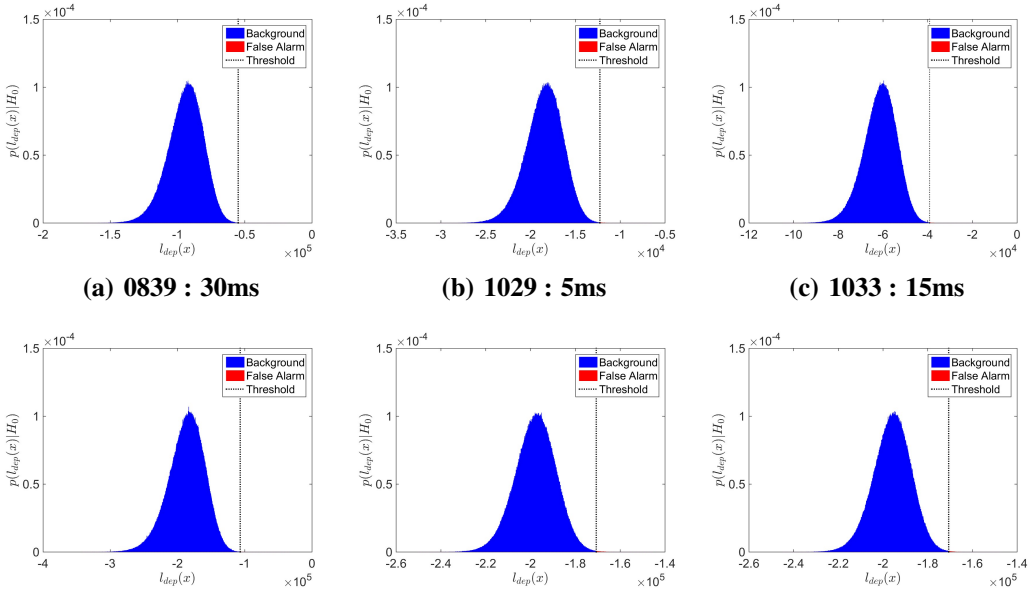
Figure D.10: Simulated Distribution Results : Red, Signal-Dependent. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure D.11: Simulated Distribution Results : Gray, Signal-Independent. Results are reflective of a goal CFAR of 10^{-3}.



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

(d) 1041 : 50ms **(e) 1431 : 30ms** **(f) 1624 : 30ms**

Figure D.12: Simulated Distribution Results : Gray, Signal-Dependent. Results are reflective of a goal CFAR of 10^{-3}.

APPENDIX E

EXPERIMENTAL DATA MODELING RESULTS - SUPPLEMENTAL TABLES

Tables E.1-E.6 contain the FAR results of all of the experimental data. Each table is representative of a single datacube. The rows labeled **Indep.**, **Static**, and **Dynamic** represent the FAR for the signal-independent detection statistic, the signal-dependent static threshold, and the signal-dependent dynamic threshold respectively. The skew and kurtosis values are also included to give reference to the Normality of the resultant distribution. These values are representative of the signal-dependent distribution, but the main body of research shows that these values do not change much between the signal-independent and signal-dependent distributions.

	Light Blue	**Blue**	**Neon**	**Pink**	**Green**	**Red**	**Yellow**	**Gray**
Indep.	1.35187	2.33896	1.09437	1.21239	2.57500	0.00000	0.91198	0.01073
Static	1.36260	2.36041	1.11583	1.22312	2.57500	0.00000	0.91198	0.01073
Dynamic	1.36260	2.36041	1.11583	1.22312	2.57500	0.00000	0.91198	0.01073
Skew	-0.03188	0.60811	-0.08756	-0.05221	0.67262	-0.16368	-0.30348	-0.16112
Kurtosis	0.114893	0.471859	0.095990	0.131324	0.448480	-0.188740	0.563834	-0.172291

Table E.1: Experimental FAR and Normality Results : 0839 Datacube. Normalized empirical FAR of experimental data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	1.74885	0.00000	1.37333	1.66302	0.00000	0.01073	1.37333	0.05365
Static	1.69521	0.00000	1.37333	1.66302	0.00000	0.01073	1.35187	0.05365
Dynamic	1.67375	0.00000	1.37333	1.66302	0.00000	0.01073	1.35187	0.05365
Skew	0.03249	-0.15733	-0.05883	-0.02063	-0.58252	-1.06584	-0.10341	-0.64961
Kurtosis	0.07974	-0.07336	0.16168	0.10934	0.16030	2.33529	0.24207	0.78764

Table E.2: Experimental FAR and Normality Results : 1029 Datacube. Normalized empirical FAR of experimental data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	2.27458	0.00000	1.95271	1.98489	0.00000	0.10729	1.81323	0.00000
Static	2.27458	0.00000	1.96344	1.98489	0.00000	0.10729	1.84541	0.00000
Dynamic	2.27458	0.00000	1.96344	1.98489	0.00000	0.10729	1.83469	0.00000
Skew	0.18212	-0.40095	0.16552	0.12143	-1.10429	-0.73785	0.13529	-0.97307
Kurtosis	0.18286	-0.04205	0.17641	0.14890	1.69844	1.36492	0.23668	1.54920

Table E.3: Experimental FAR and Normality Results : 1033 Datacube. Normalized empirical FAR of experimental data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	2.81104	0.00000	2.67156	2.36041	0.00000	0.05365	2.98270	0.00000
Static	2.82177	0.00000	2.68229	2.36041	0.00000	0.05365	3.01489	0.00000
Dynamic	2.82177	0.00000	2.68229	2.36041	0.00000	0.05365	3.00416	0.00000
Skew	0.25483	-0.67390	0.24374	0.16086	-1.30533	-0.56145	0.21119	-1.15814
Kurtosis	0.29003	0.72215	0.36170	0.23365	2.32462	0.70208	0.46654	1.96287

Table E.4: Experimental FAR and Normality Results : 1041 Datacube. Normalized empirical FAR of experimental data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	16.53363	0.07510	15.22467	14.74186	5.69718	7.53187	13.67967	7.90739
Static	16.65165	0.07510	15.29977	14.84915	5.76156	7.57478	13.74405	7.97176
Dynamic	16.61946	0.07510	15.29977	14.81696	5.75083	7.57478	13.73332	7.95030
Skew	0.29563	-1.26739	0.25295	0.24712	-0.26044	0.01821	0.28336	-0.11744
Kurtosis	0.32169	3.48600	0.58021	0.35743	0.38913	0.55027	0.93429	0.27848

Table E.5: Experimental FAR and Normality Results : 1431 Datacube. Normalized empirical FAR of experimental data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

	Light Blue	Blue	Neon	Pink	Green	Red	Yellow	Gray
Indep.	15.28904	5.64353	16.09373	15.47144	6.48041	12.53165	15.40706	5.55770
Static	15.29977	5.67572	16.10446	15.50363	6.53405	12.57457	15.44998	5.56843
Dynamic	15.29977	5.66499	16.10446	15.48217	6.53405	12.57457	15.43925	5.55770
Skew	0.09755	0.00202	0.10010	0.05325	0.03395	-0.06049	0.14841	0.00882
Kurtosis	0.92490	0.60950	1.18301	1.00155	0.57956	1.03764	1.33992	0.53732

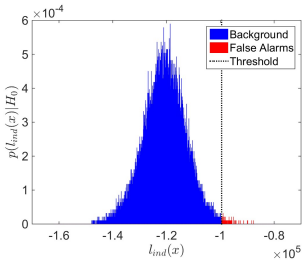
Table E.6: Experimental FAR and Normality Results : 1624 Datacube. Normalized empirical FAR of experimental data. Results shown have been multiplied by a factor of 10^3 so that a value of 1 represents a perfect match with the desired FAR of 10^{-3}.

APPENDIX F

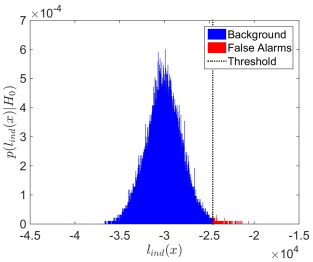
EXPERIMENTAL DATA MODELING RESULTS - SUPPLEMENTAL FIGURES

Figures F.1 - F.24 display the signal-independent and signal-dependent detection statistics for all results along with a false alarm map for each result. These results are calculated in an identical manner to Figures 5.19 - 5.20 in the main body of research.

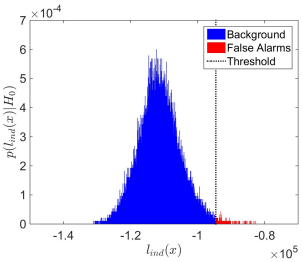
Light Blue Target
Signal-Independent Detection Statistic



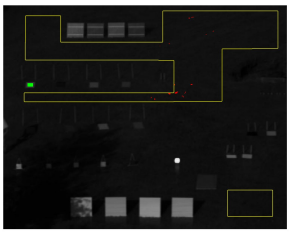
(a) 0839 : 30ms



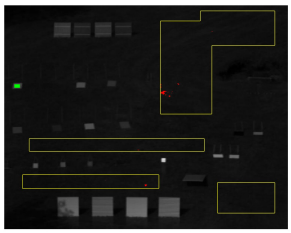
(b) 1029 : 5ms



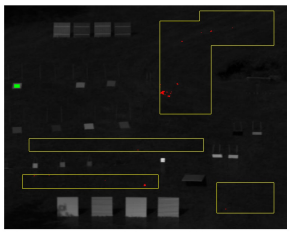
(c) 1033 : 15ms



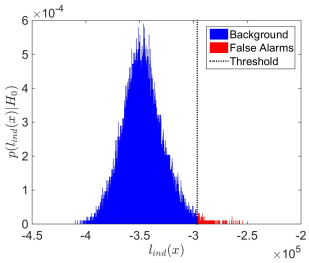
(d) 0839 : 30ms



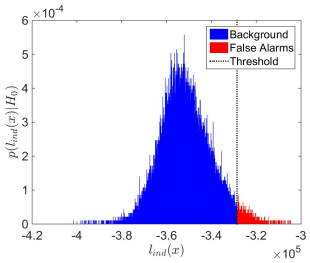
(e) 1029 : 5ms



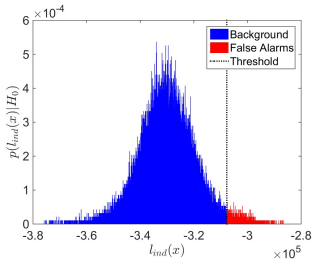
(f) 1033 : 15ms



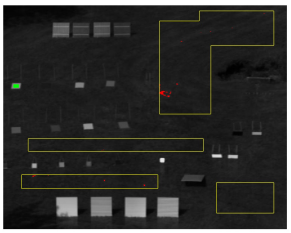
(g) 1041 : 50ms



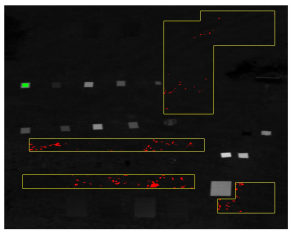
(h) 1431 : 30ms



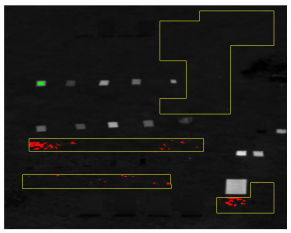
(i) 1624 : 30ms



(j) 1041 : 50ms



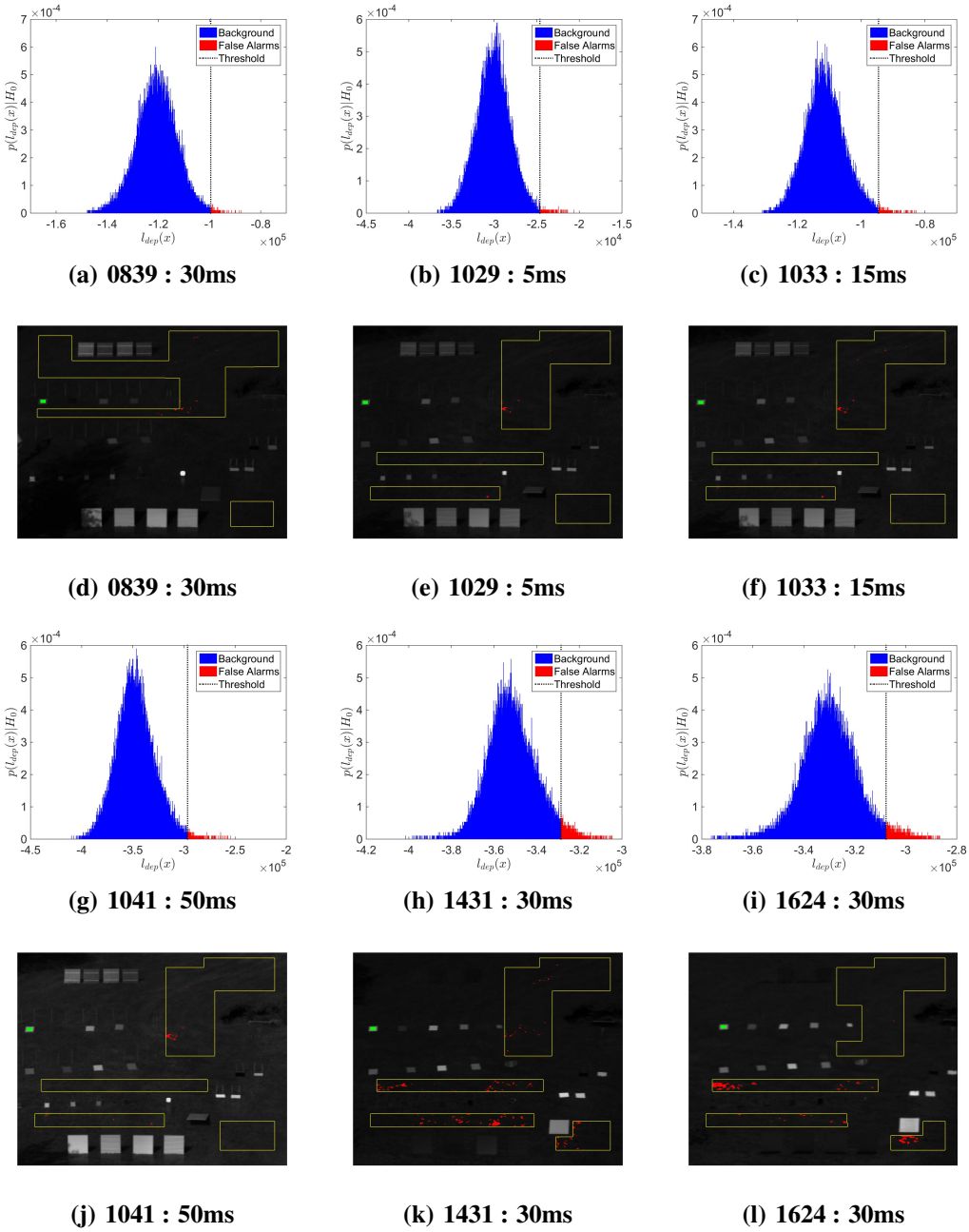
(k) 1431 : 30ms



(l) 1624 : 30ms

Figure F.1: Experimental Results : Light Blue, Signal-Independent.

Light Blue Target
Signal-Dependent, Static Threshold Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

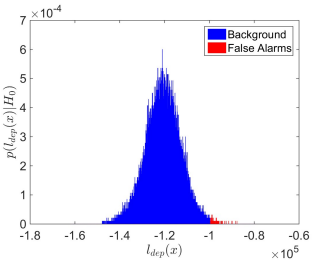
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

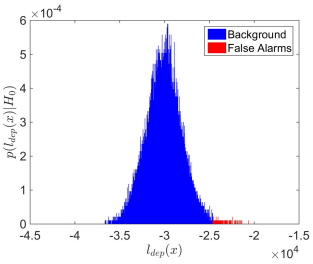
(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.2: Experimental Results : Light Blue, Signal-Dependent, Static.

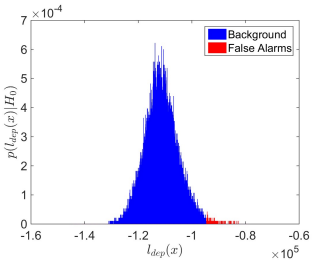
Light Blue Target
Signal-Dependent, Dynamic Threshold Detection Statistic



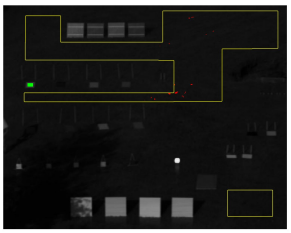
(a) 0839 : 30ms



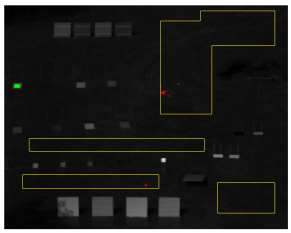
(b) 1029 : 5ms



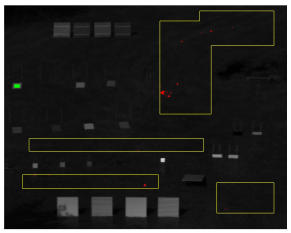
(c) 1033 : 15ms



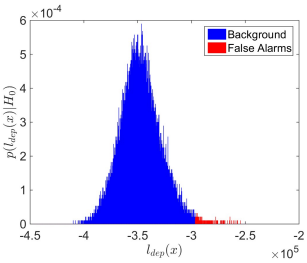
(d) 0839 : 30ms



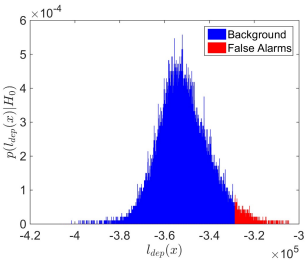
(e) 1029 : 5ms



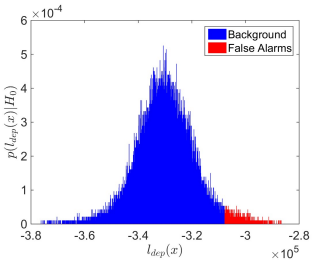
(f) 1033 : 15ms



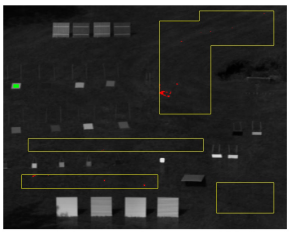
(g) 1041 : 50ms



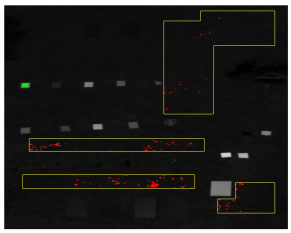
(h) 1431 : 30ms



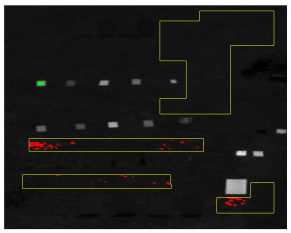
(i) 1624 : 30ms



(j) 1041 : 50ms



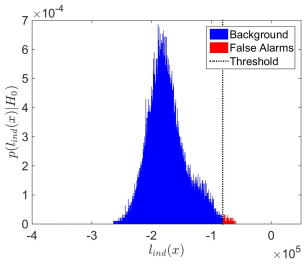
(k) 1431 : 30ms



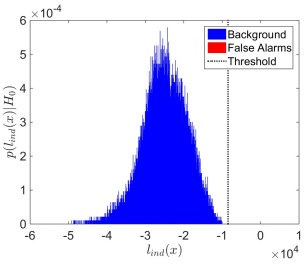
(l) 1624 : 30ms

Figure F.3: Experimental Results : Light Blue, Signal-Dependent, Dynamic.

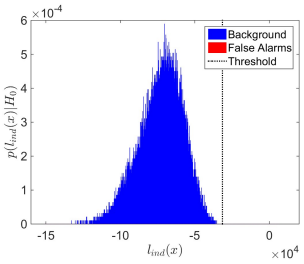
Blue Target
Signal-Independent Detection Statistic



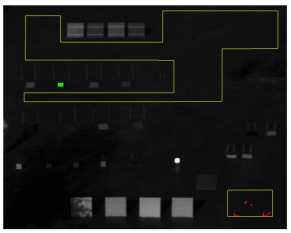
(a) 0839 : 30ms



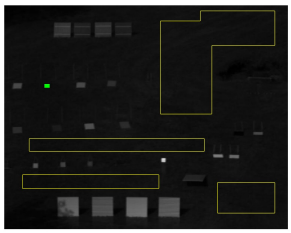
(b) 1029 : 5ms



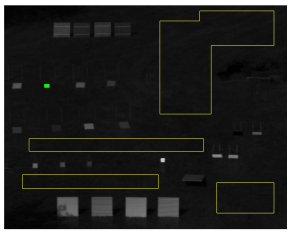
(c) 1033 : 15ms



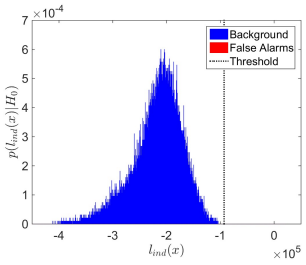
(d) 0839 : 30ms



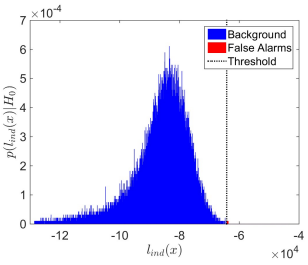
(e) 1029 : 5ms



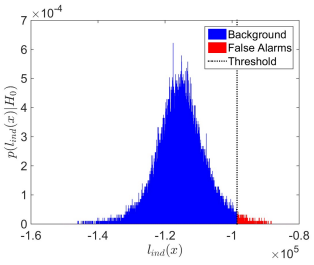
(f) 1033 : 15ms



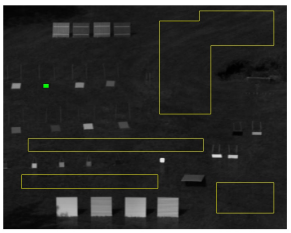
(g) 1041 : 50ms



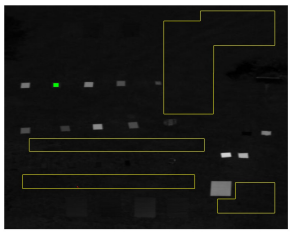
(h) 1431 : 30ms



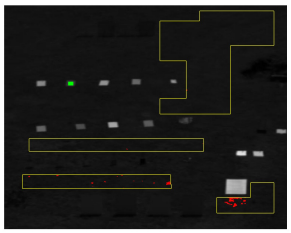
(i) 1624 : 30ms



(j) 1041 : 50ms



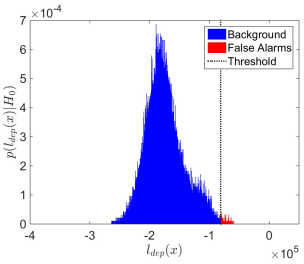
(k) 1431 : 30ms



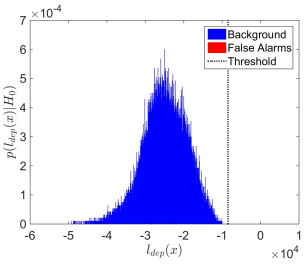
(l) 1624 : 30ms

Figure F.4: Experimental Results : Blue, Signal-Independent.

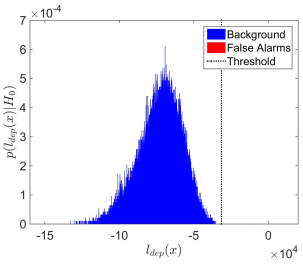
Blue Target
Signal-Dependent, Static Threshold Detection Statistic



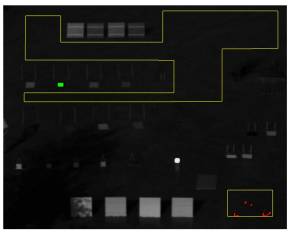
(a) 0839 : 30ms



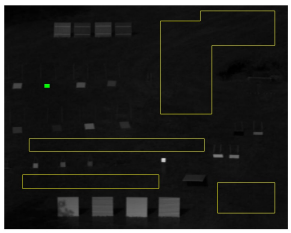
(b) 1029 : 5ms



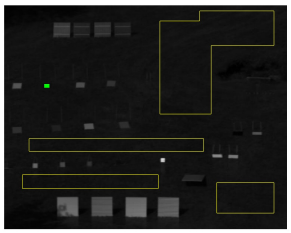
(c) 1033 : 15ms



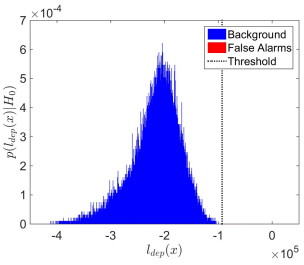
(d) 0839 : 30ms



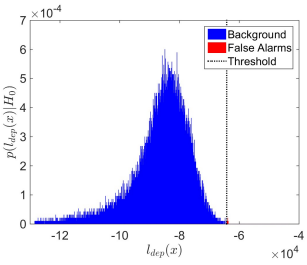
(e) 1029 : 5ms



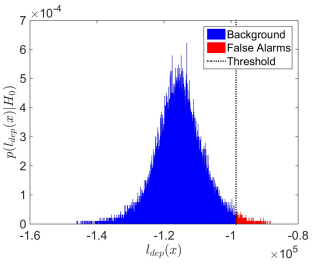
(f) 1033 : 15ms



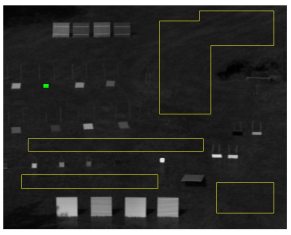
(g) 1041 : 50ms



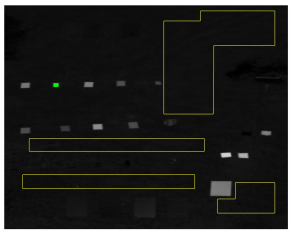
(h) 1431 : 30ms



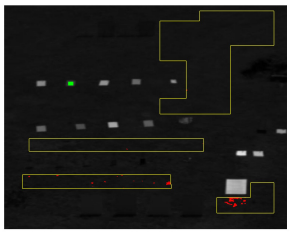
(i) 1624 : 30ms



(j) 1041 : 50ms



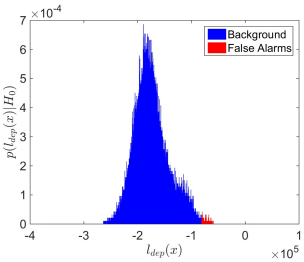
(k) 1431 : 30ms



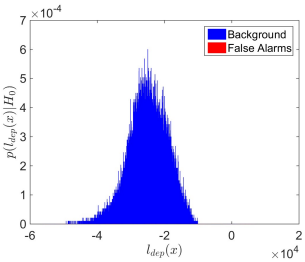
(l) 1624 : 30ms

Figure F.5: Experimental Results : Blue, Signal-Dependent, Static.

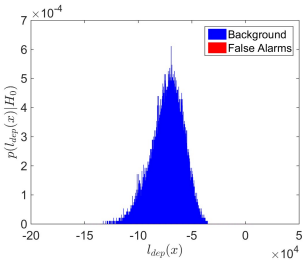
Blue Target
Signal-Dependent, Dynamic Threshold Detection Statistic



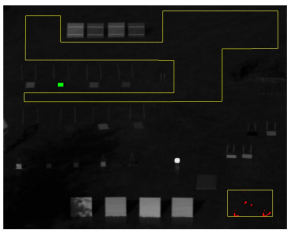
(a) 0839 : 30ms



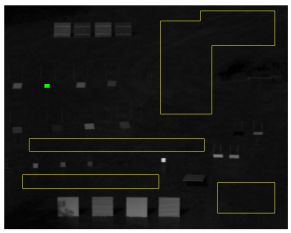
(b) 1029 : 5ms



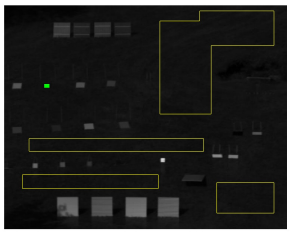
(c) 1033 : 15ms



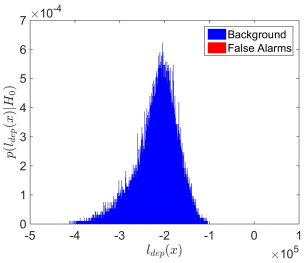
(d) 0839 : 30ms



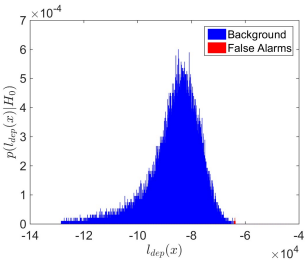
(e) 1029 : 5ms



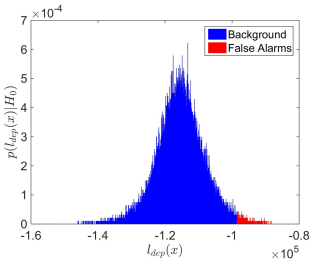
(f) 1033 : 15ms



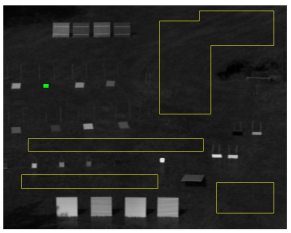
(g) 1041 : 50ms



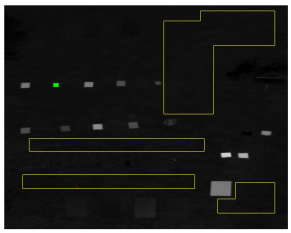
(h) 1431 : 30ms



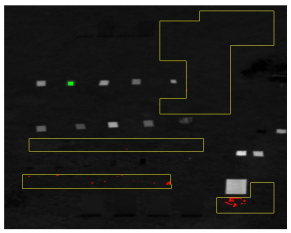
(i) 1624 : 30ms



(j) 1041 : 50ms



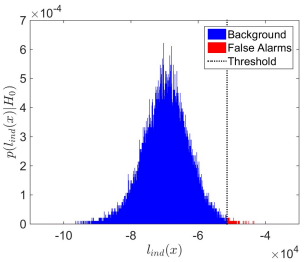
(k) 1431 : 30ms



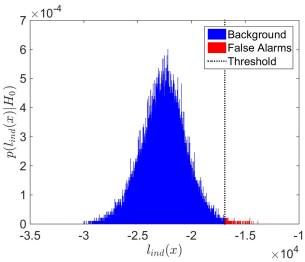
(l) 1624 : 30ms

Figure F.6: Experimental Results : Blue, Signal-Dependent, Dynamic.

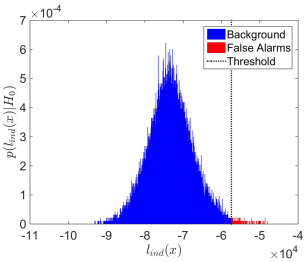
Neon Target
Signal-Independent Detection Statistic



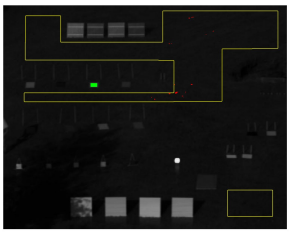
(a) 0839 : 30ms



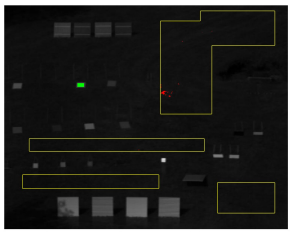
(b) 1029 : 5ms



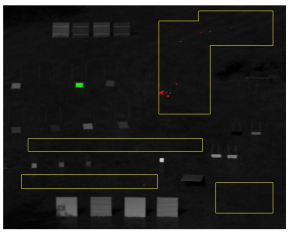
(c) 1033 : 15ms



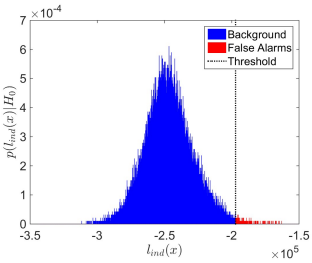
(d) 0839 : 30ms



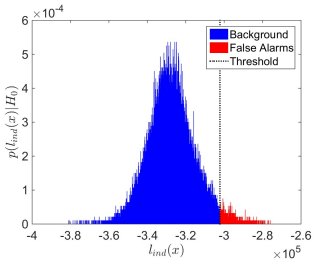
(e) 1029 : 5ms



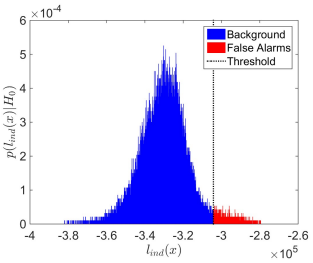
(f) 1033 : 15ms



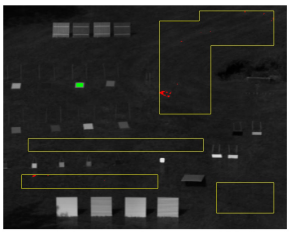
(g) 1041 : 50ms



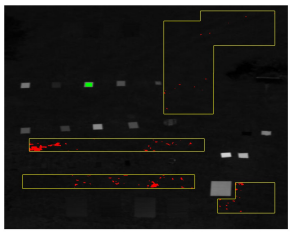
(h) 1431 : 30ms



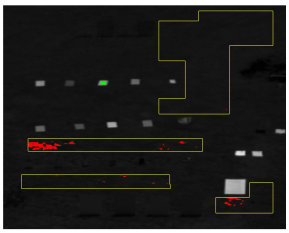
(i) 1624 : 30ms



(j) 1041 : 50ms



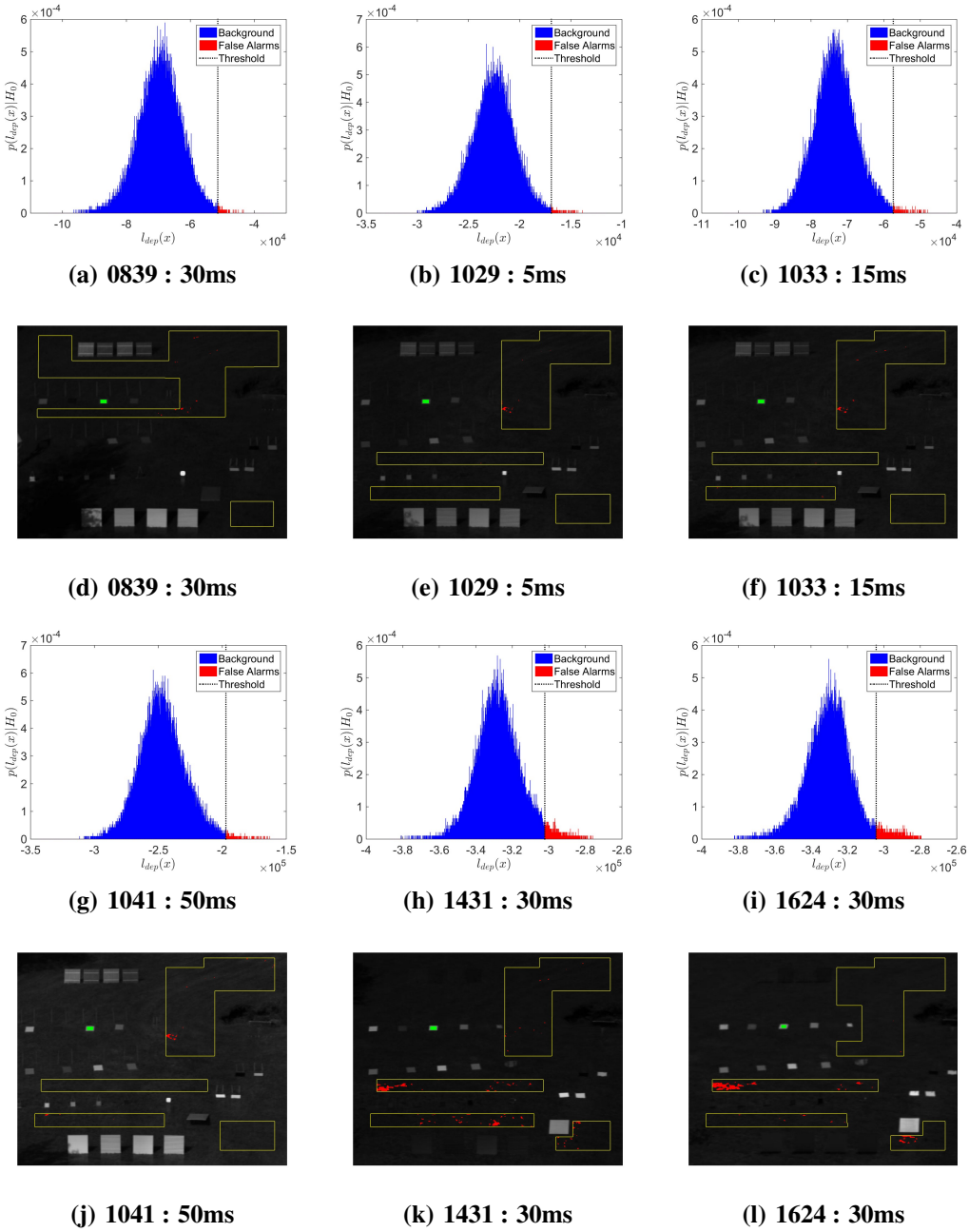
(k) 1431 : 30ms



(l) 1624 : 30ms

Figure F.7: Experimental Results : Neon Green, Signal-Independent.

Neon Target
Signal-Dependent, Static Threshold Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

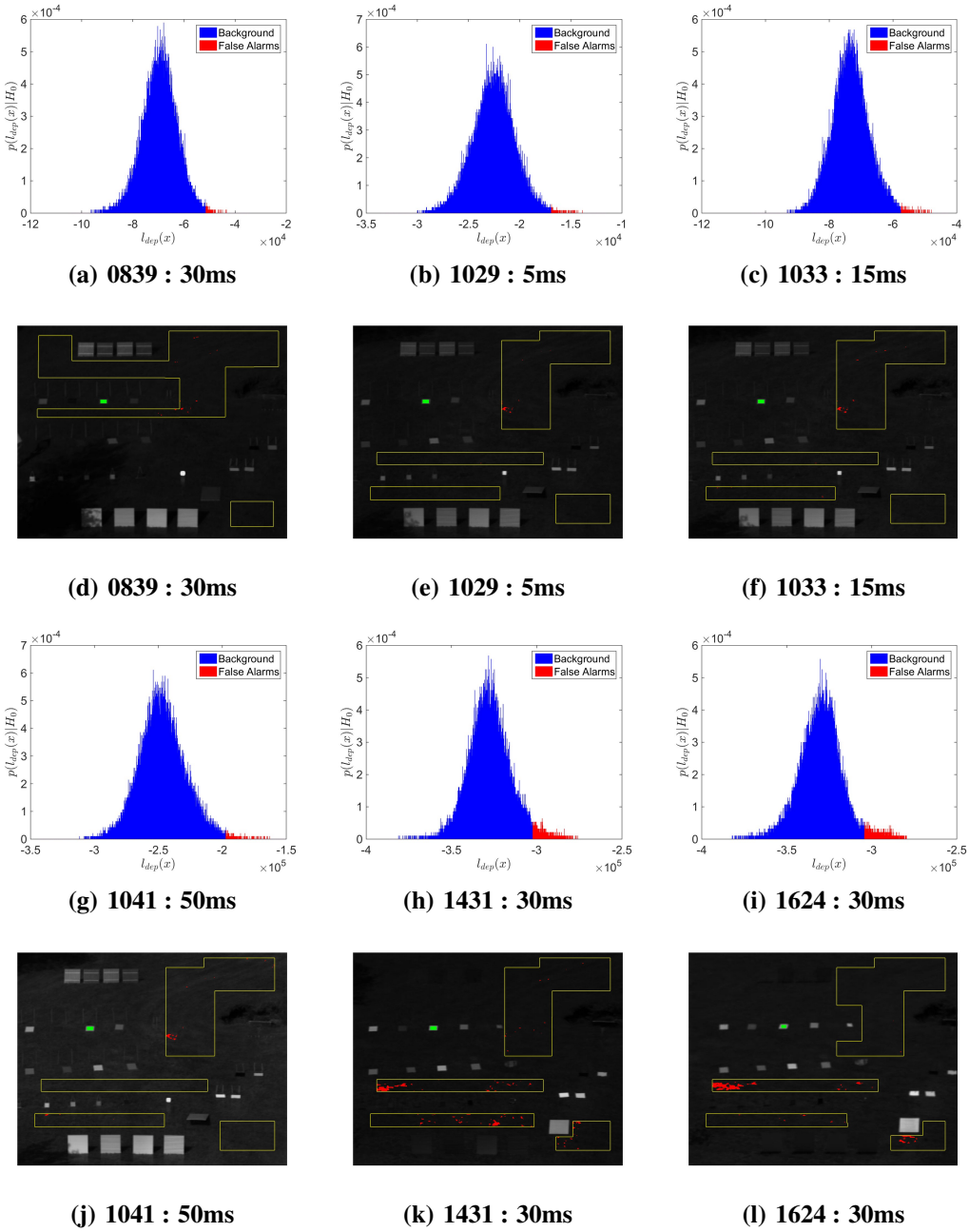
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.8: Experimental Results : Neon Green, Signal-Dependent, Static.

Neon Target
Signal-Dependent, Dynamic Threshold Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

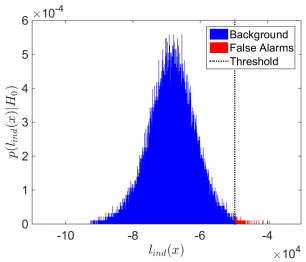
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

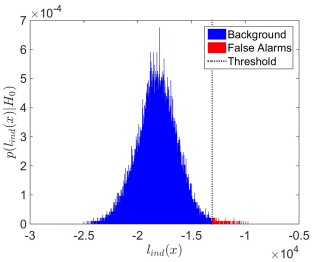
(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.9: Experimental Results : Neon Green, Signal-Dependent, Dynamic.

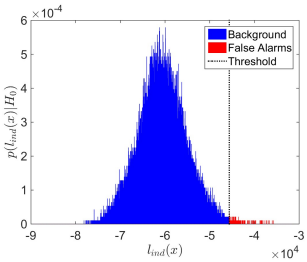
Pink Target
Signal-Independent Detection Statistic



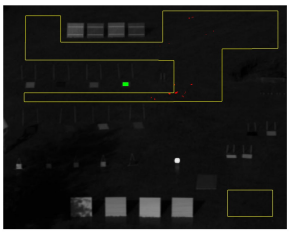
(a) 0839 : 30ms



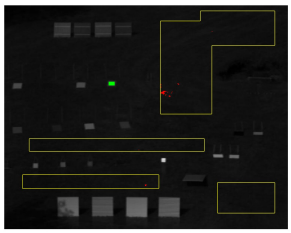
(b) 1029 : 5ms



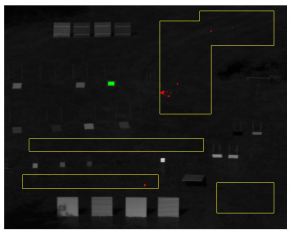
(c) 1033 : 15ms



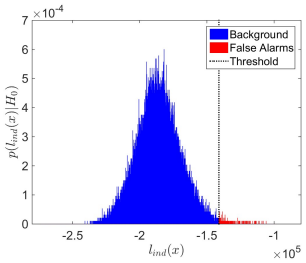
(d) 0839 : 30ms



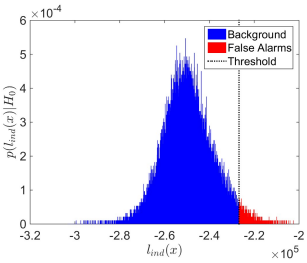
(e) 1029 : 5ms



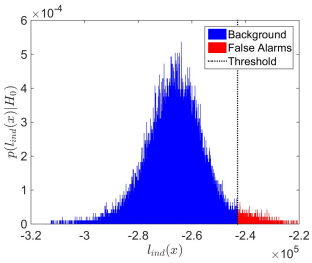
(f) 1033 : 15ms



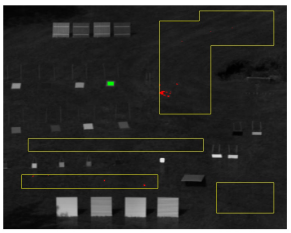
(g) 1041 : 50ms



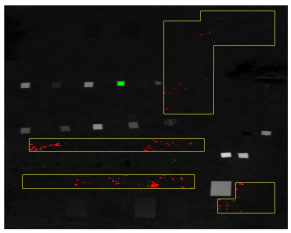
(h) 1431 : 30ms



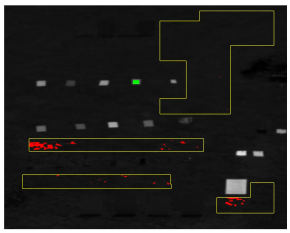
(i) 1624 : 30ms



(j) 1041 : 50ms



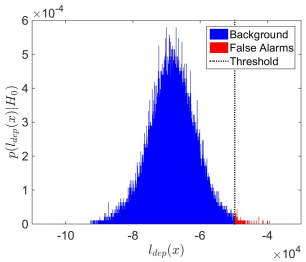
(k) 1431 : 30ms



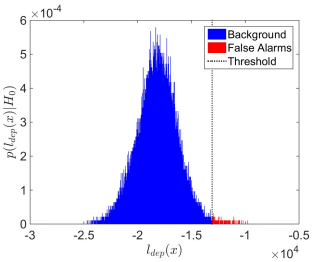
(l) 1624 : 30ms

Figure F.10: Experimental Results : Pink, Signal-Independent.

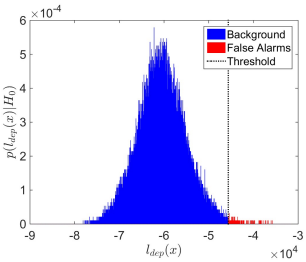
Pink Target
Signal-Dependent, Static Threshold Detection Statistic



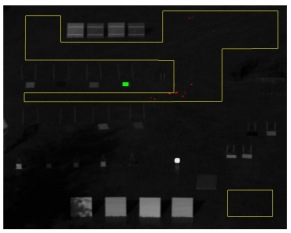
(a) 0839 : 30ms



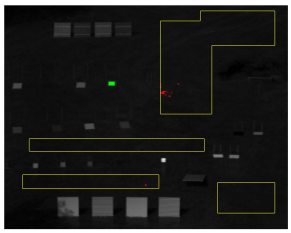
(b) 1029 : 5ms



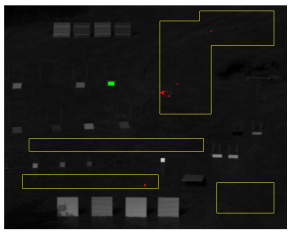
(c) 1033 : 15ms



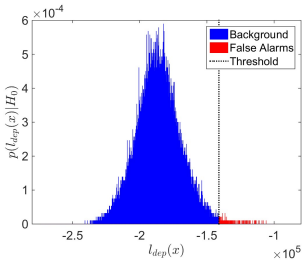
(d) 0839 : 30ms



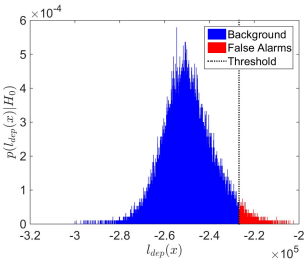
(e) 1029 : 5ms



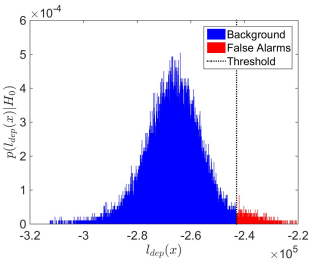
(f) 1033 : 15ms



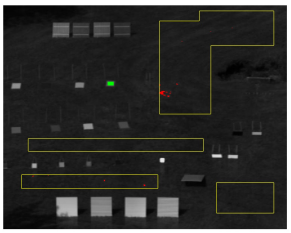
(g) 1041 : 50ms



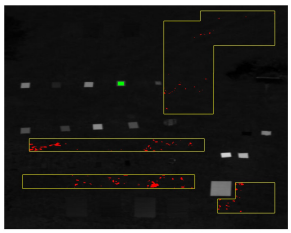
(h) 1431 : 30ms



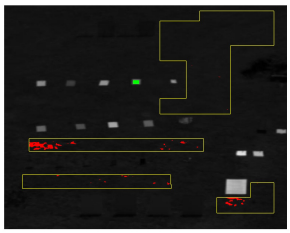
(i) 1624 : 30ms



(j) 1041 : 50ms



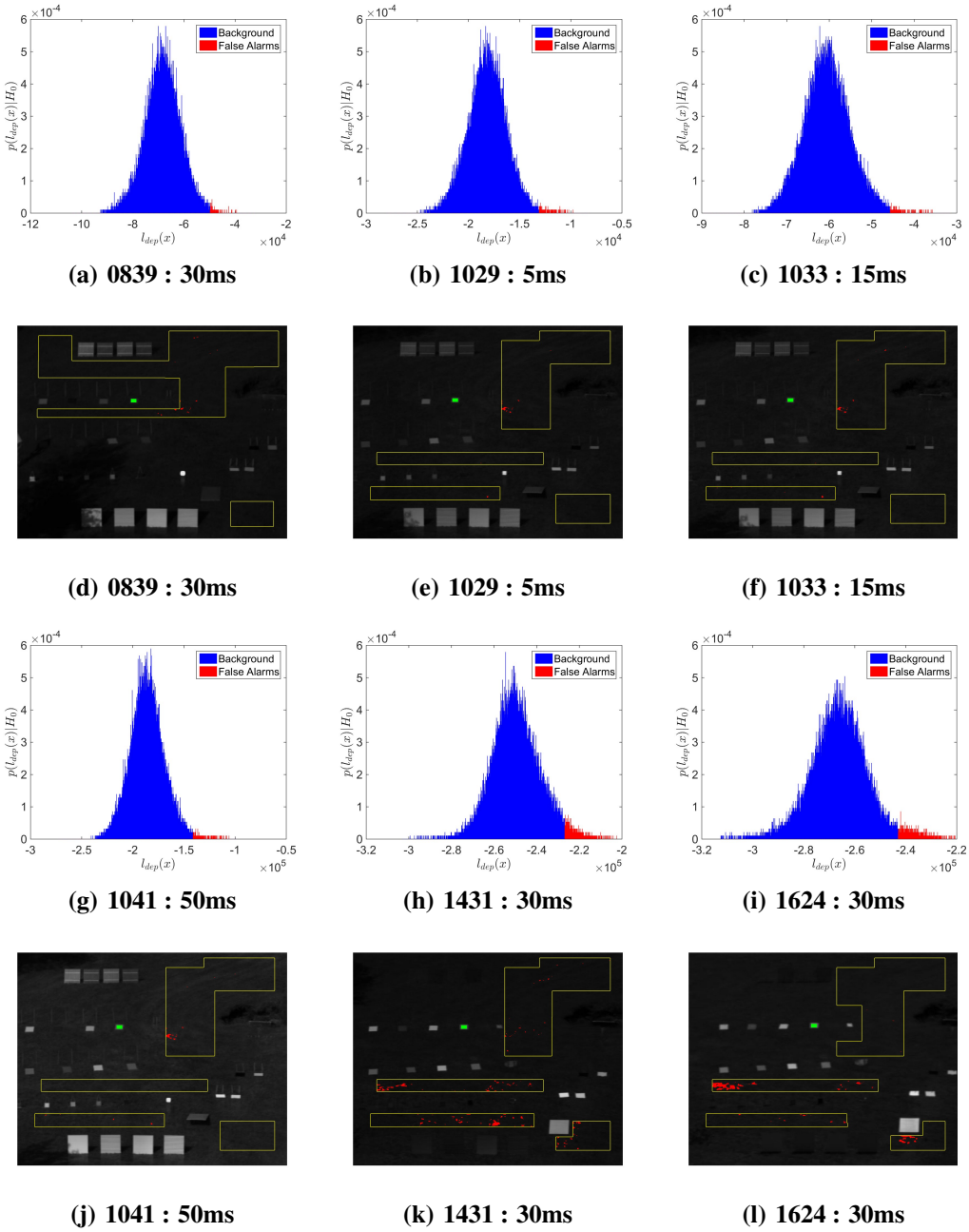
(k) 1431 : 30ms



(l) 1624 : 30ms

Figure F.11: Experimental Results : Pink, Signal-Dependent, Static.

Pink Target
Signal-Dependent, Dynamic Threshold Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

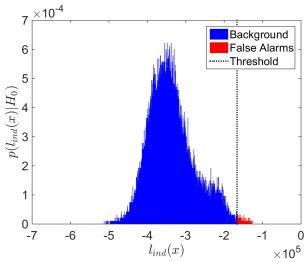
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

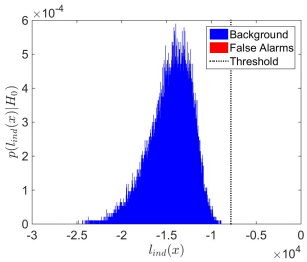
(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.12: Experimental Results : Pink, Signal-Dependent, Dynamic.

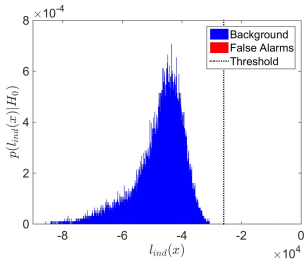
Green Target
Signal-Independent Detection Statistic



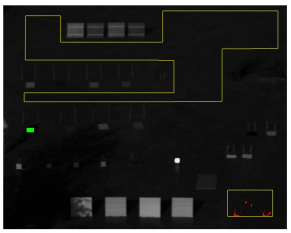
(a) 0839 : 30ms



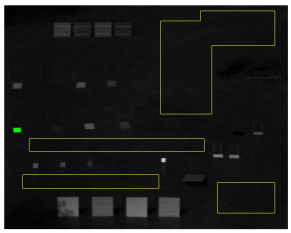
(b) 1029 : 5ms



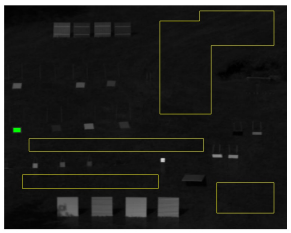
(c) 1033 : 15ms



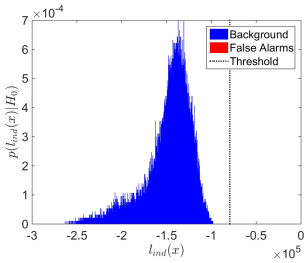
(d) 0839 : 30ms



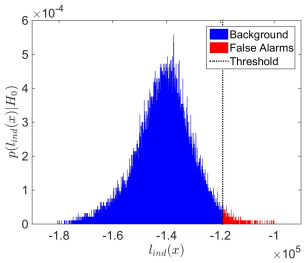
(e) 1029 : 5ms



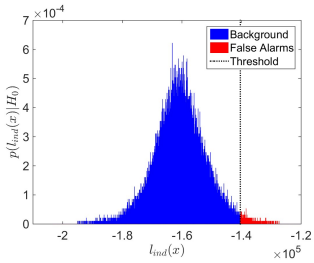
(f) 1033 : 15ms



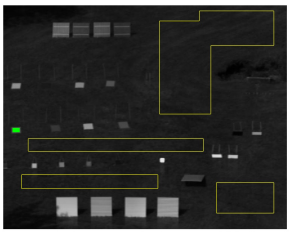
(g) 1041 : 50ms



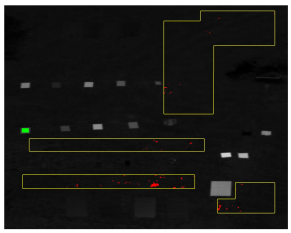
(h) 1431 : 30ms



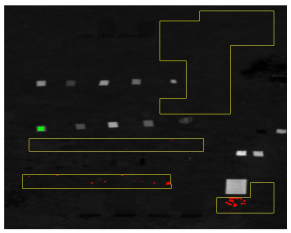
(i) 1624 : 30ms



(j) 1041 : 50ms



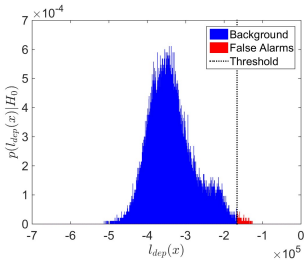
(k) 1431 : 30ms



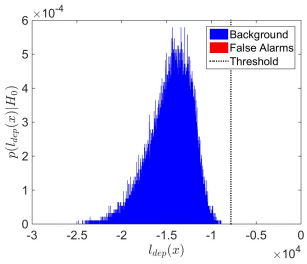
(l) 1624 : 30ms

Figure F.13: Experimental Results : Green, Signal-Independent.

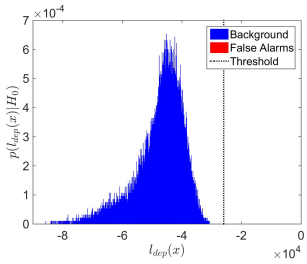
Green Target
Signal-Dependent, Static Threshold Detection Statistic



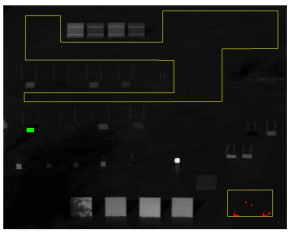
(a) 0839 : 30ms



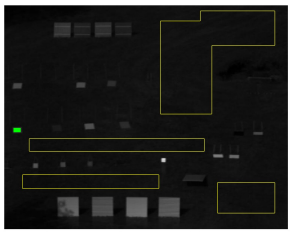
(b) 1029 : 5ms



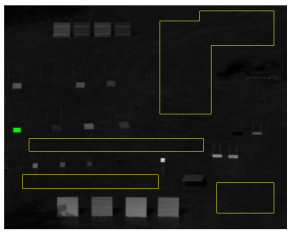
(c) 1033 : 15ms



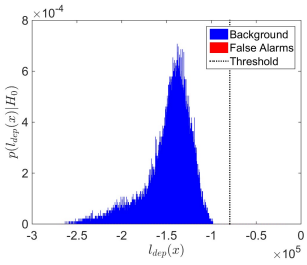
(d) 0839 : 30ms



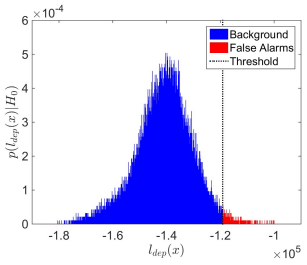
(e) 1029 : 5ms



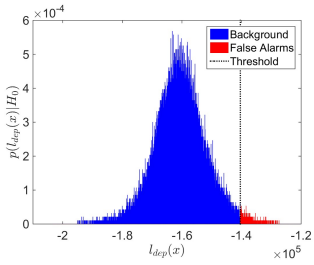
(f) 1033 : 15ms



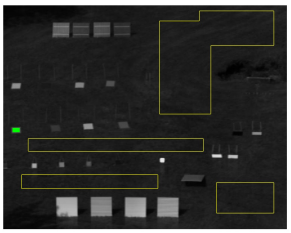
(g) 1041 : 50ms



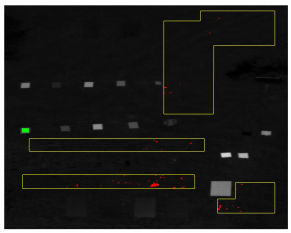
(h) 1431 : 30ms



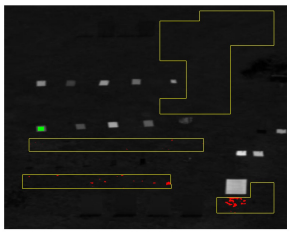
(i) 1624 : 30ms



(j) 1041 : 50ms



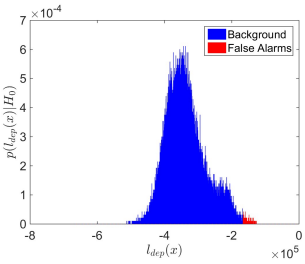
(k) 1431 : 30ms



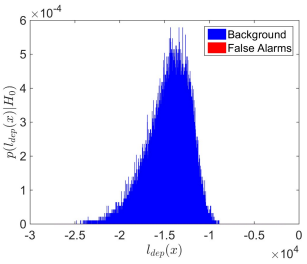
(l) 1624 : 30ms

Figure F.14: Experimental Results : Green, Signal-Dependent, Static.

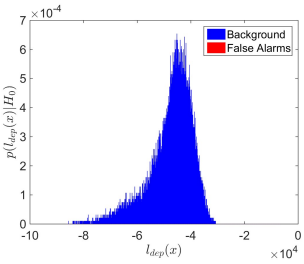
Green Target
Signal-Dependent, Dynamic Threshold Detection Statistic



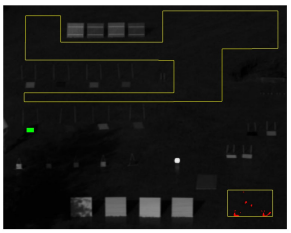
(a) 0839 : 30ms



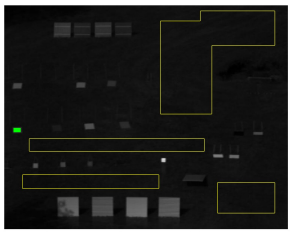
(b) 1029 : 5ms



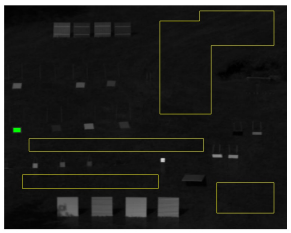
(c) 1033 : 15ms



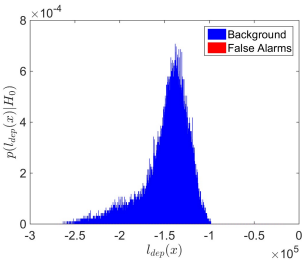
(d) 0839 : 30ms



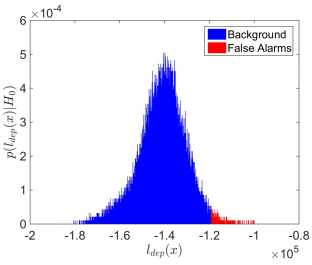
(e) 1029 : 5ms



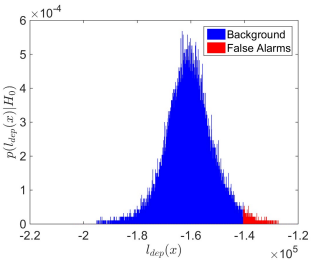
(f) 1033 : 15ms



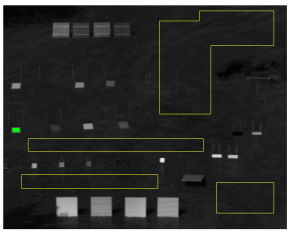
(g) 1041 : 50ms



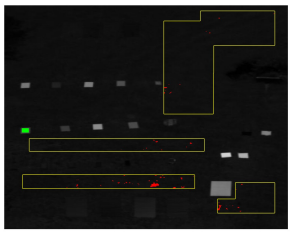
(h) 1431 : 30ms



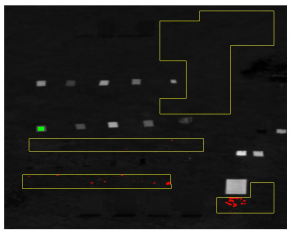
(i) 1624 : 30ms



(j) 1041 : 50ms



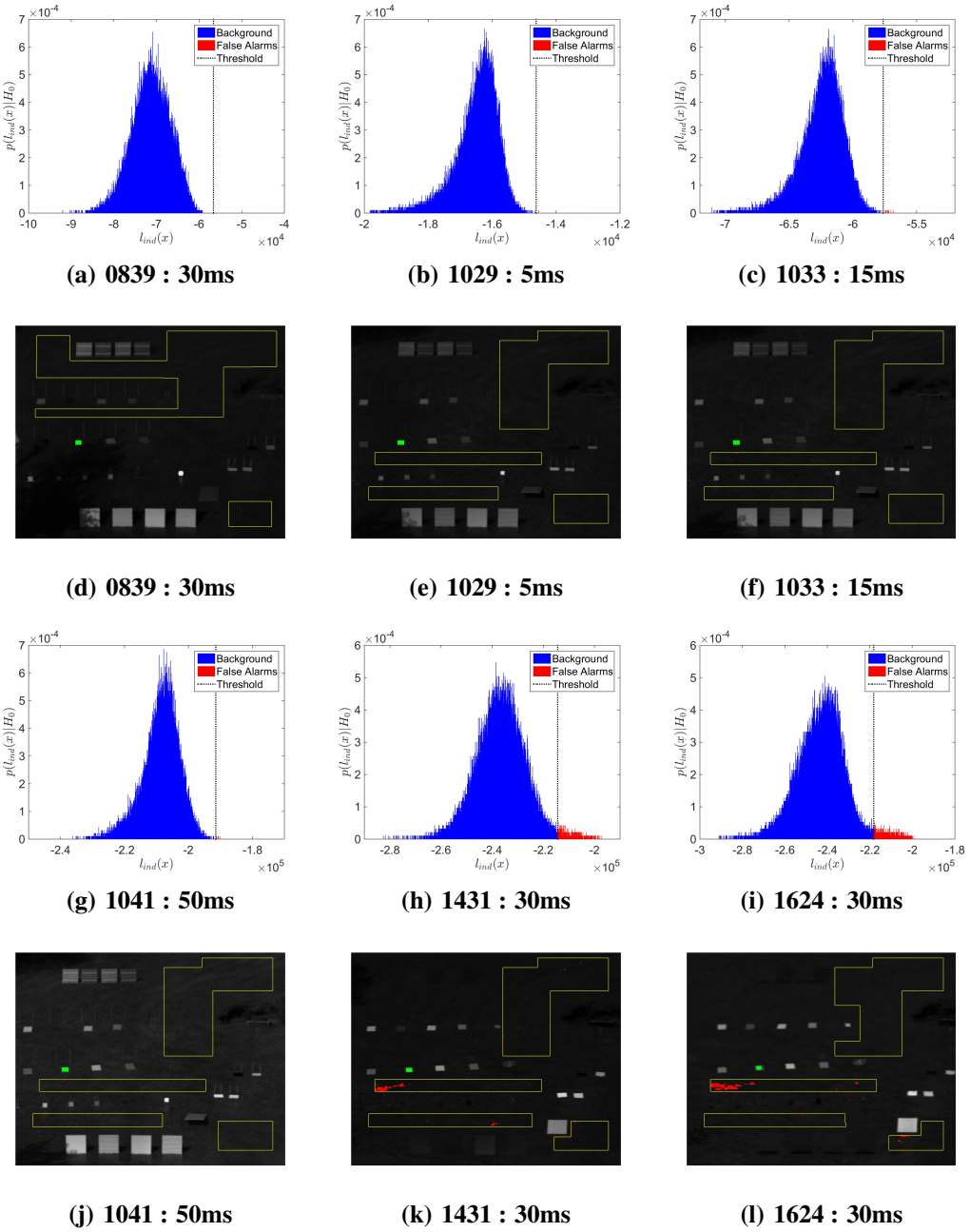
(k) 1431 : 30ms



(l) 1624 : 30ms

Figure F.15: Experimental Results : Green, Signal-Dependent, Dynamic.

Red Target
Signal-Independent Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

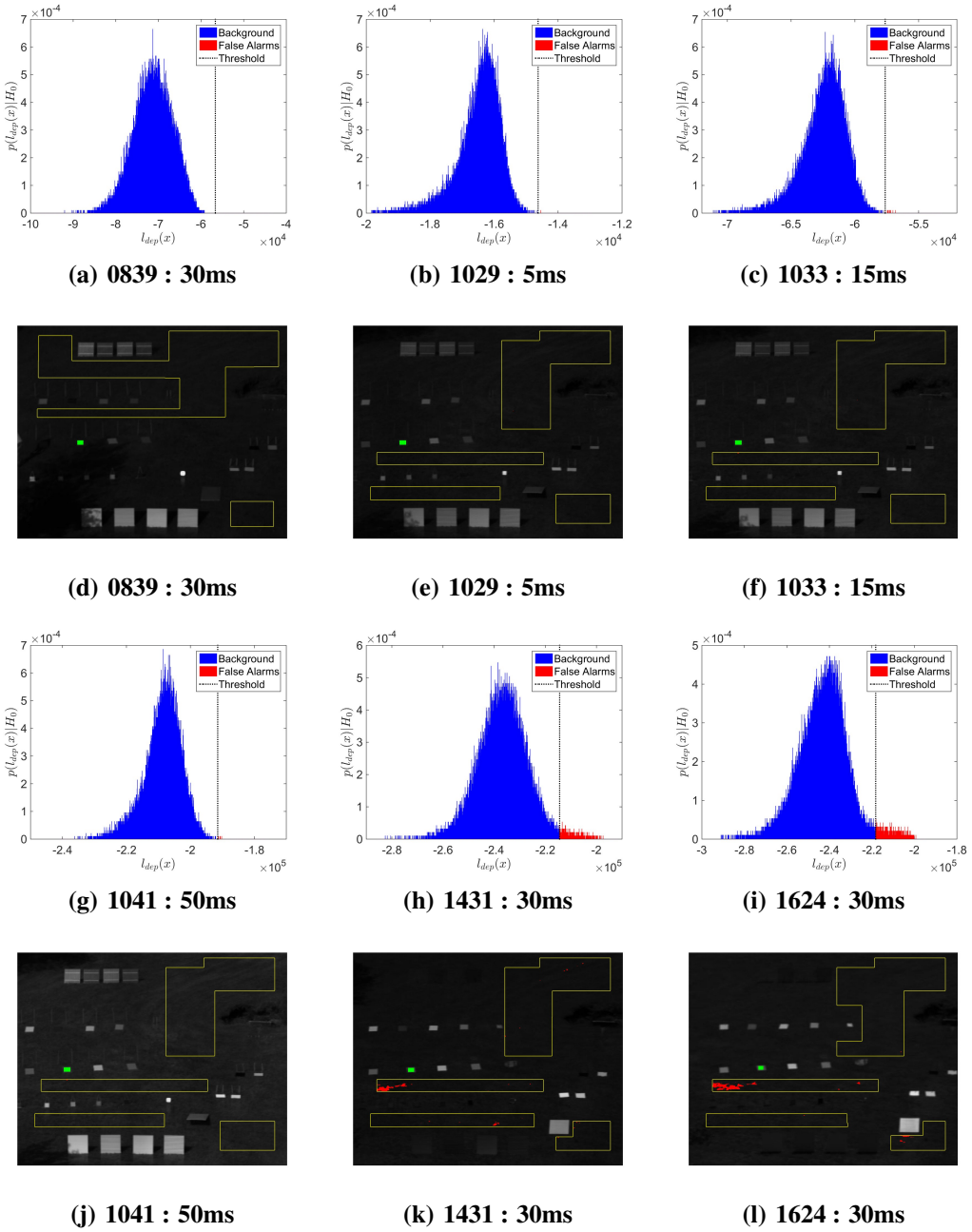
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.16: Experimental Results : Red, Signal-Independent.

Red Target
Signal-Dependent, Static Threshold Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

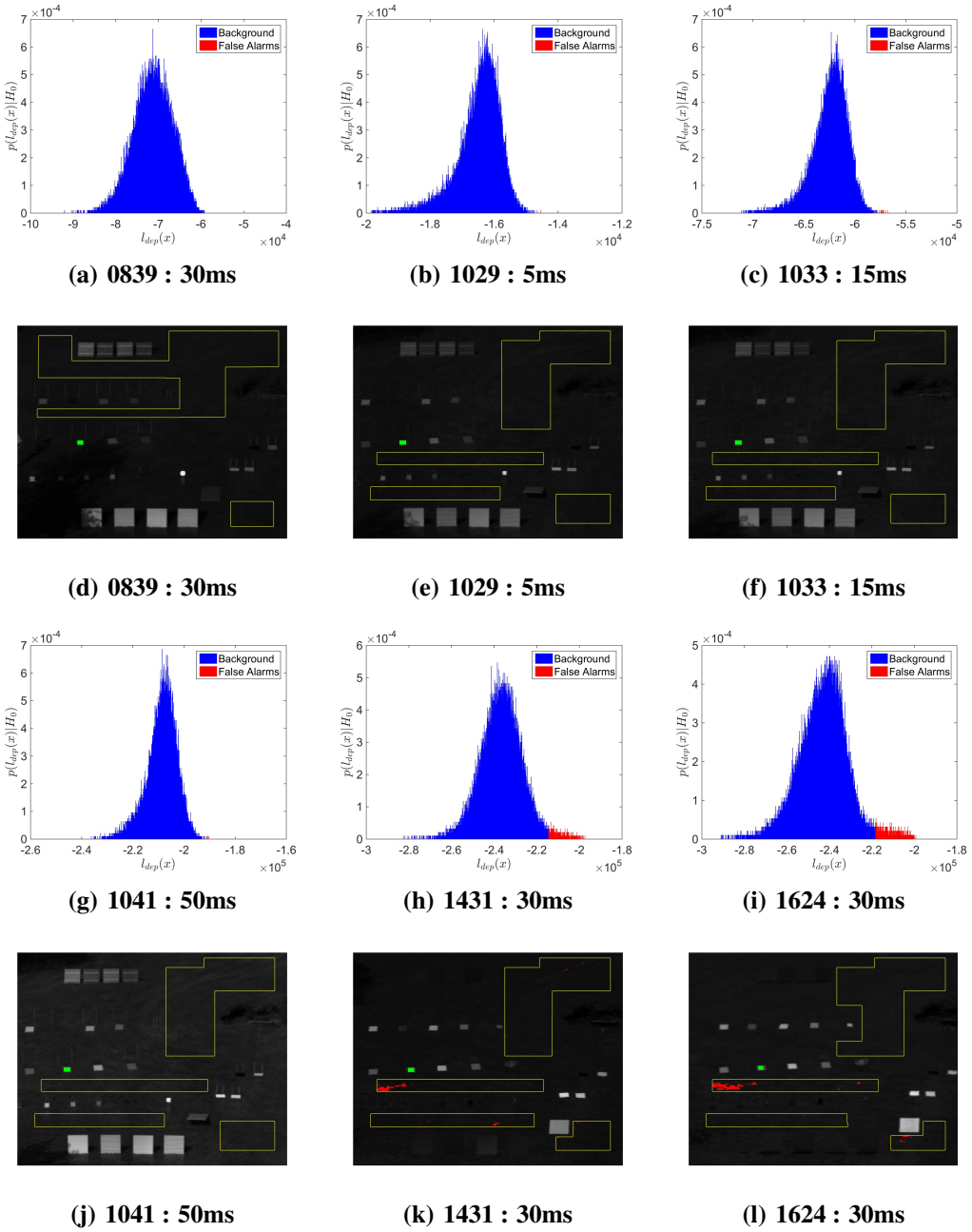
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.17: Experimental Results : Red, Signal-Dependent, Static.

Red Target
Signal-Dependent, Dynamic Threshold Detection Statistic



(a) 0839 : 30ms **(b) 1029 : 5ms** **(c) 1033 : 15ms**

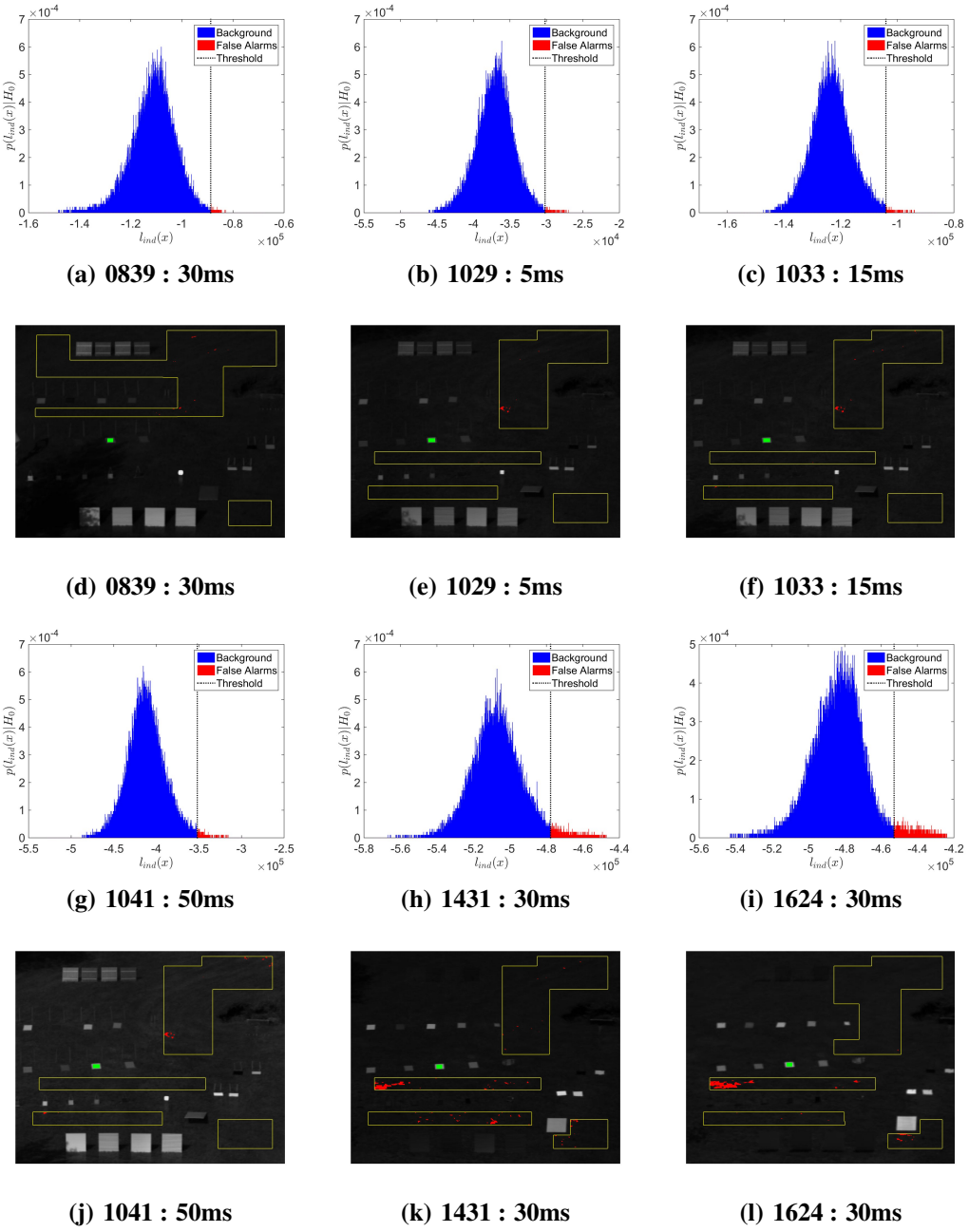
(d) 0839 : 30ms **(e) 1029 : 5ms** **(f) 1033 : 15ms**

(g) 1041 : 50ms **(h) 1431 : 30ms** **(i) 1624 : 30ms**

(j) 1041 : 50ms **(k) 1431 : 30ms** **(l) 1624 : 30ms**

Figure F.18: Experimental Results : Red, Signal-Dependent, Dynamic.

Yellow Target
Signal-Independent Detection Statistic



(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms

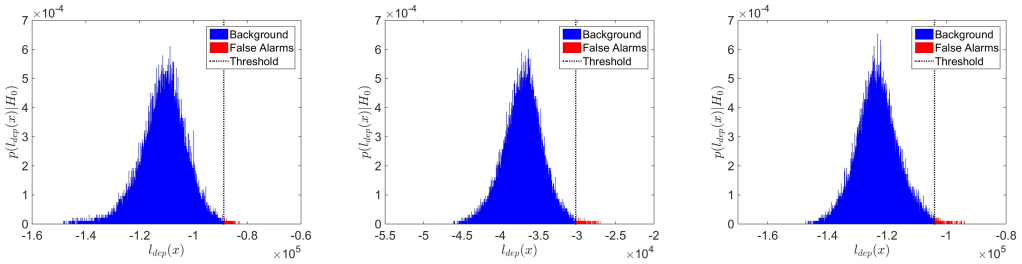
(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms

(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms

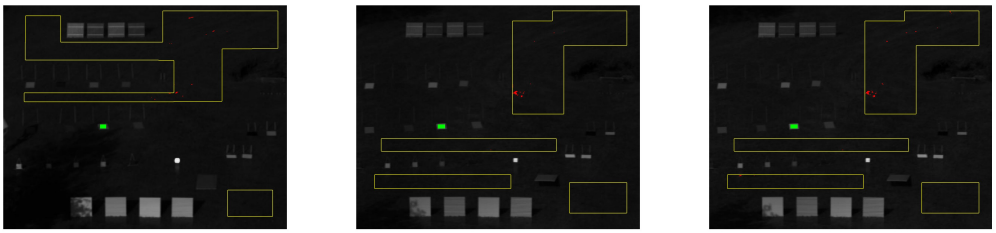
(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.19: Experimental Results : Yellow, Signal-Independent.

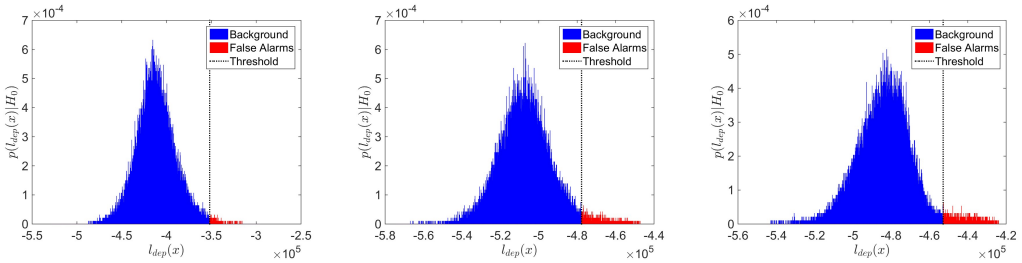
Yellow Target
Signal-Dependent, Static Threshold Detection Statistic



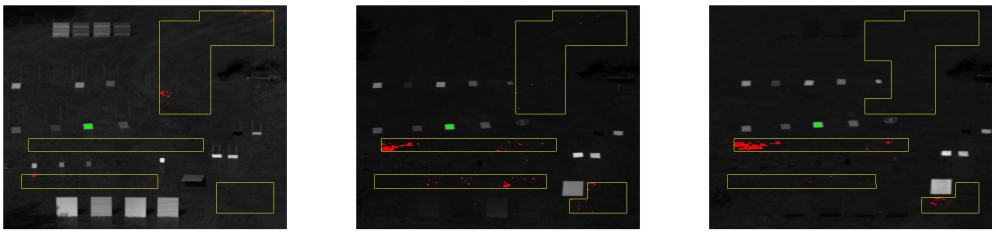
(a) 0839 : 30ms (b) 1029 : 5ms (c) 1033 : 15ms



(d) 0839 : 30ms (e) 1029 : 5ms (f) 1033 : 15ms



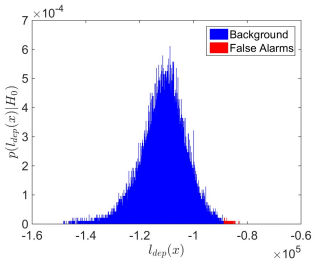
(g) 1041 : 50ms (h) 1431 : 30ms (i) 1624 : 30ms



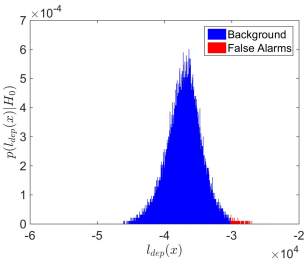
(j) 1041 : 50ms (k) 1431 : 30ms (l) 1624 : 30ms

Figure F.20: Experimental Results : Yellow, Signal-Dependent, Static.

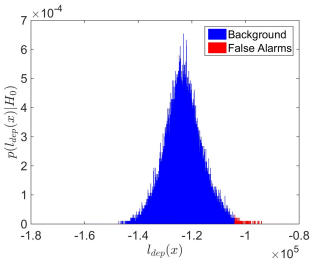
Yellow Target
Signal-Dependent, Dynamic Threshold Detection Statistic



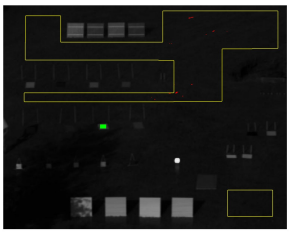
(a) 0839 : 30ms



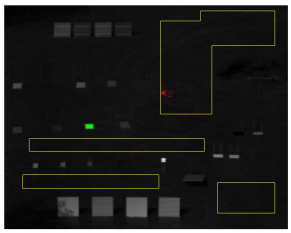
(b) 1029 : 5ms



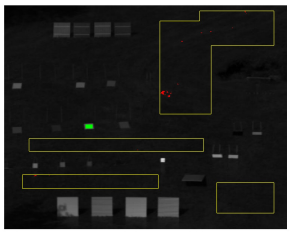
(c) 1033 : 15ms



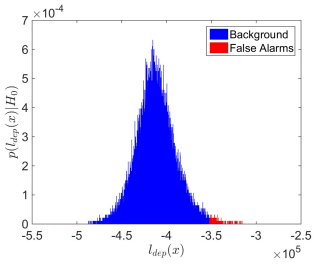
(d) 0839 : 30ms



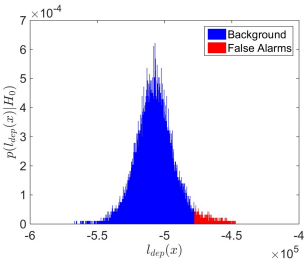
(e) 1029 : 5ms



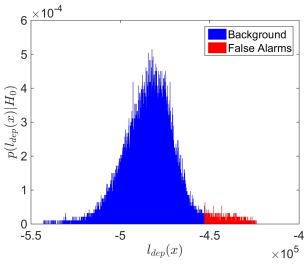
(f) 1033 : 15ms



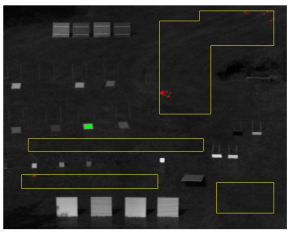
(g) 1041 : 50ms



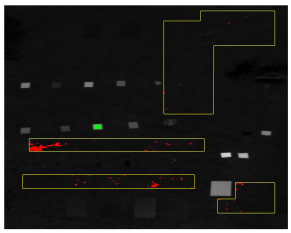
(h) 1431 : 30ms



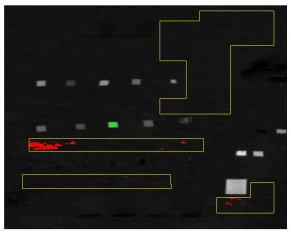
(i) 1624 : 30ms



(j) 1041 : 50ms



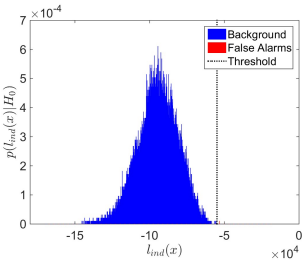
(k) 1431 : 30ms



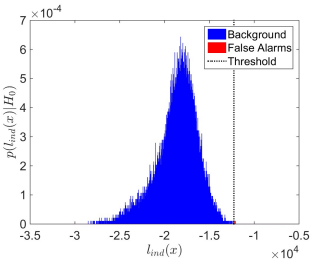
(l) 1624 : 30ms

Figure F.21: Experimental Results : Yellow, Signal-Dependent, Dynamic.

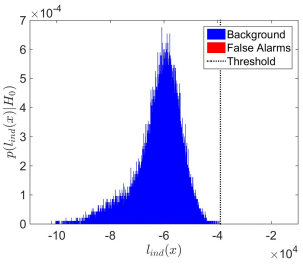
Gray Target
Signal-Independent Detection Statistic



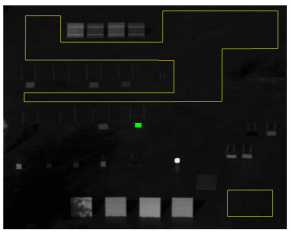
(a) 0839 : 30ms



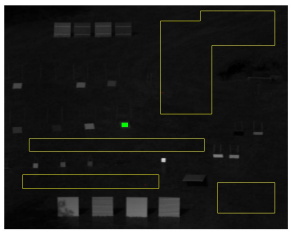
(b) 1029 : 5ms



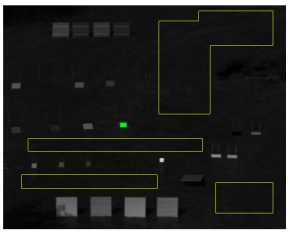
(c) 1033 : 15ms



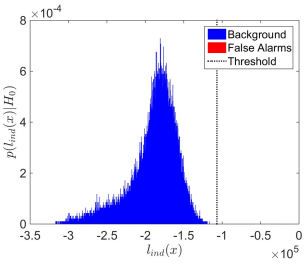
(d) 0839 : 30ms



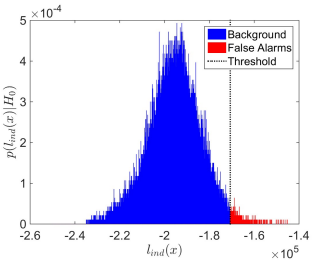
(e) 1029 : 5ms



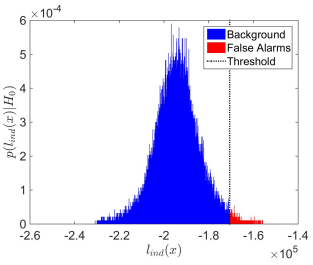
(f) 1033 : 15ms



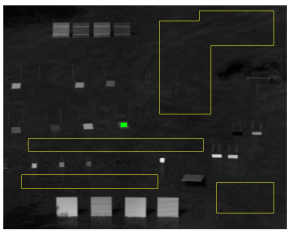
(g) 1041 : 50ms



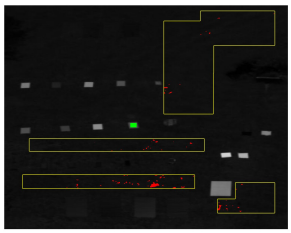
(h) 1431 : 30ms



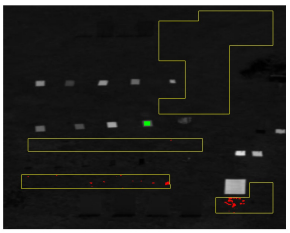
(i) 1624 : 30ms



(j) 1041 : 50ms



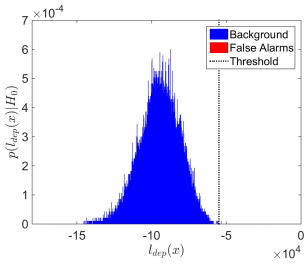
(k) 1431 : 30ms



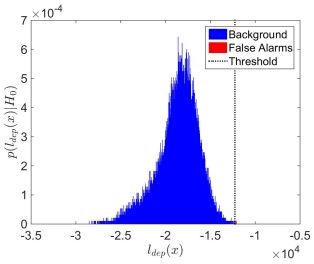
(l) 1624 : 30ms

Figure F.22: Experimental Results : Gray, Signal-Independent.

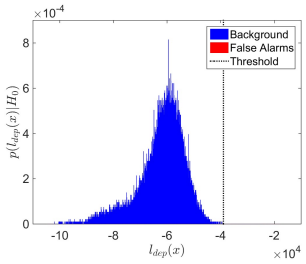
Gray Target
Signal-Dependent, Static Threshold Detection Statistic



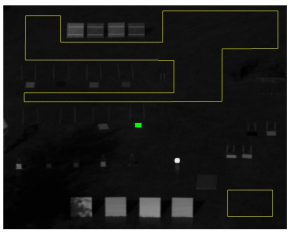
(a) 0839 : 30ms



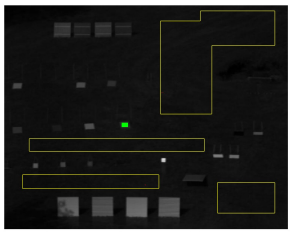
(b) 1029 : 5ms



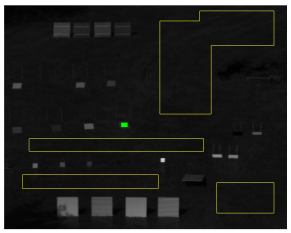
(c) 1033 : 15ms



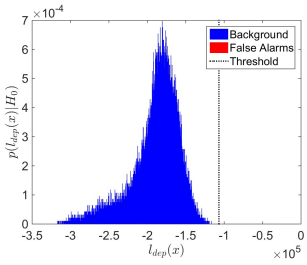
(d) 0839 : 30ms



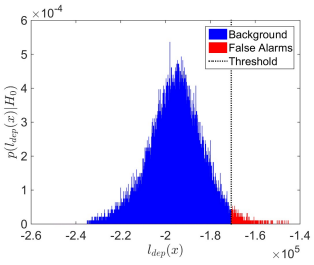
(e) 1029 : 5ms



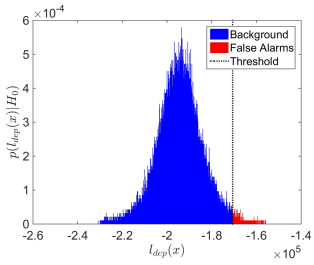
(f) 1033 : 15ms



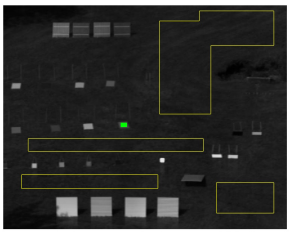
(g) 1041 : 50ms



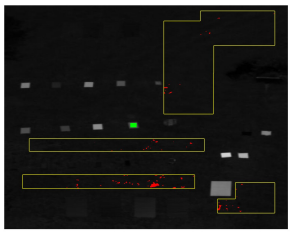
(h) 1431 : 30ms



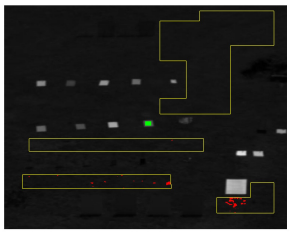
(i) 1624 : 30ms



(j) 1041 : 50ms



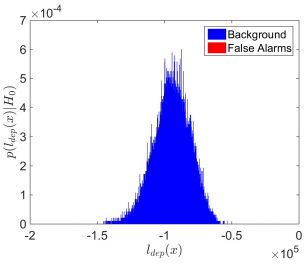
(k) 1431 : 30ms



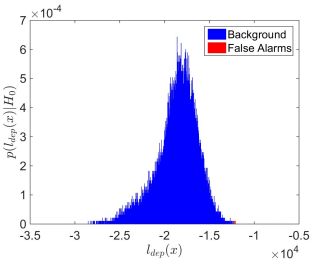
(l) 1624 : 30ms

Figure F.23: Experimental Results : Gray, Signal-Dependent, Static.

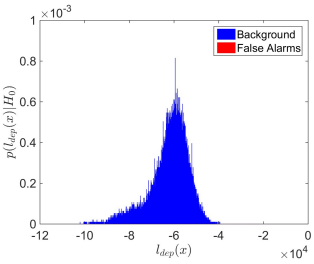
Gray Target
Signal-Dependent, Dynamic Threshold Detection Statistic



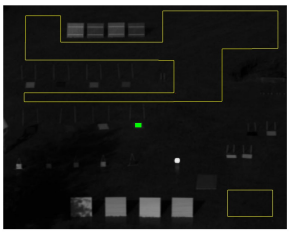
(a) 0839 : 30ms



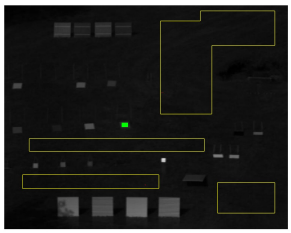
(b) 1029 : 5ms



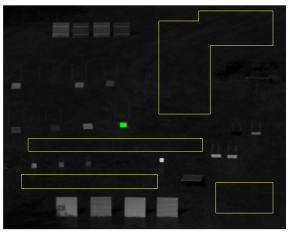
(c) 1033 : 15ms



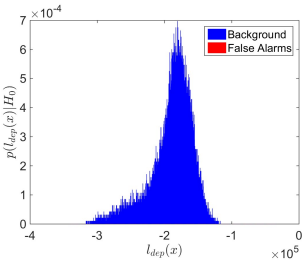
(d) 0839 : 30ms



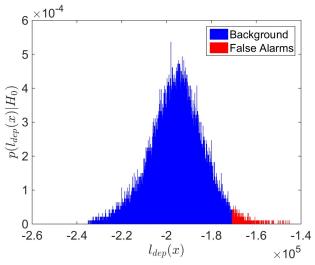
(e) 1029 : 5ms



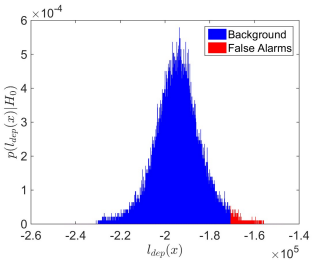
(f) 1033 : 15ms



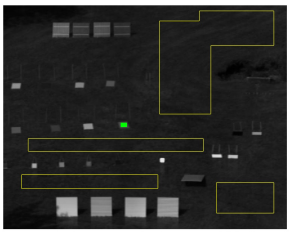
(g) 1041 : 50ms



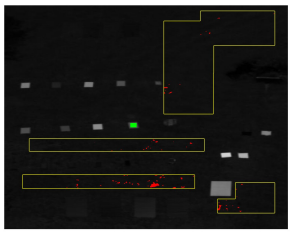
(h) 1431 : 30ms



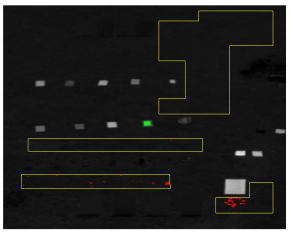
(i) 1624 : 30ms



(j) 1041 : 50ms



(k) 1431 : 30ms



(l) 1624 : 30ms

Figure F.24: Experimental Results : Gray, Signal-Dependent, Dynamic.